Bayesian Comparative Statics*

Teddy Mekonnen †   René Leal Vizcaíno ‡

First Draft: November 2016

Current Draft: March 2018

Abstract

An agent chooses an action after observing a signal about an unknown state. A more informative signal leads to a more dispersed distribution of the agent's posterior beliefs. For a subclass of supermodular utility functions, we show that a more informative signal also leads to a more dispersed distribution of actions along with monotone changes to the average optimal action. Furthermore, outside the subclass of utility functions we consider, we provide an example of a more informative signal that does not lead to a more dispersed distribution of actions. We extend our results to Bayesian games with strategic complementarities and show that a more informative signal for any one player leads to a more dispersed distribution of equilibrium actions for all players. We apply these results to study informational externalities and to compare the demand for information in covert and overt information acquisition games.

Keywords: Supermodularity, comparative statics, informativeness, decision-making under uncertainty, games of incomplete information, informational externalities

JEL Codes: C44, C61, D42, D81

*This paper was previously circulated under the title, "Ex-ante Comparative Statics: Responsiveness to Information Quality." We are indebted to Asher Wolinsky, Eddie Dekel, Bruno Strulovici, and Jeff Ely for their unwavering guidance and encouragement. Our work has also greatly benefited from conversations with Laura Doval, Jay Lu, Alessandro Pavan, Rob Porter, Phil Reny, Marciano Siniscalchi, Jean Tirole, and Zenon Zabinski.

†California Institute of Technology, SISL. Contact: mekonnen@caltech.edu

‡Northwestern University, Department of Economics. Contact: renelealv@u.northwestern.edu

1 Introduction

An agent faces a decision-making problem under uncertainty. After observing a signal that is informative about an unknown state of the world, the agent chooses an action. Since the agent's choice depends on the signal realization she receives, from the perspective of an uninformed observer, the agent's optimal action is an endogenously determined random variable.

In this paper, we characterize how the quality of the agent's signal affects the induced distribution of her optimal action. We consider a setting in which the agent prefers to take higher actions for higher states of the world. When the agent's signal becomes more informative about the unknown state, her posterior beliefs become more dispersed. We characterize under what conditions a more dispersed distribution of posterior beliefs leads to (i) a more dispersed distribution of actions, and (ii) a higher or lower mean of the optimal action. Our characterization elucidates the mechanism behind the familiar result that a mean-preserving spread in posterior beliefs leads to a mean-preserving spread of actions in the case of quadratic payoffs and Gaussian signals.[1] Moreover, our characterization extends the same intuition to allow for changes in the mean as well as more general payoff functions and signal structures.

In various settings, the agent's action not only affects her own payoff but also the payoff of other economic agents. For example, in games of incomplete information, an agent's action affects the payoff of all other players. Since the quality of the agent's signal affects the distribution of her optimal action, it also affects the (ex-post) payoff distribution of such third-party agents, thereby generating *informational externalities*. We apply our comparative statics results on the distribution of actions to study these informational externalities.

To concretely motivate our question, consider a monopolist facing a linear demand curve $P(q) = 1 - q$ and a quadratic cost function $c(\theta, q) = (1 - \theta)q + q^2/2$, where q is the quantity produced and $\theta \in [0, 1]$ is a cost parameter. Higher values of θ correspond to lower marginal costs. Consequently, the monopolist would like to produce more as θ increases.

However, the cost parameter is an unobserved random variable that is uniformly distributed on the unit interval. The monopolist instead observes a signal such that with probability $\rho \in [0, 1]$, the signal realization s matches the realized state of the world ($s = \theta$), and with probability $1 - \rho$, the signal realization s is uniformly drawn from the unit interval and independent from the state variable. The quality of the signal is increasing in ρ: When $\rho = 0$, the signal is uninformative; when $\rho = 1$, the signal is fully revealing.

From an "interim" perspective, a monopolist that observes a signal realization s when the

[1] See Mas-Colell et al. (1995), Section 6.D.2, for a definition of mean-preserving spreads.

signal quality is ρ optimally produces

$$q^M(s;\rho) = \frac{E[\theta]}{3} + \rho\left(\frac{s - E[\theta]}{3}\right).$$

The first term on the right-hand-side is the quantity the monopolist produces based only on the prior. The second term reflects how the monopolist adjusts her production decision based on what she learns from observing a signal realization $s \in [0,1]$.

From the perspective of an uninformed third party, the optimal quantity is a random variable whose distribution is given by

$$H(z;\rho) = \mathbb{P}(\{s : q^M(s;\rho) \leq z\}),$$

the probability that the monopolist optimally produces at most z units given a signal of quality ρ. Our goal in this paper is to characterize how $H(\cdot;\rho)$ changes when ρ increases. Will the optimal quantity produced increase or decrease on average when ρ increases? Will the quantities produced become more dispersed? In Section 2.5, we tackle these questions for more general utility functions and information structures (signals). However, in this example, we can answer these questions by using the closed-form solution of q^M.

Suppose the quality of the monopolist's information structure increases from ρ' to $\rho'' > \rho'$. "Good news" ($s > E[\theta]$) from ρ'' provides a stronger evidence of high values of θ than "good news" from ρ'. As a result, the monopolist produces more when she observes "good news" from ρ'' than when she observes "good news" from ρ'. Symmetrically, "bad news" ($s < E[\theta]$) from ρ'' provides a stronger evidence of low values of θ than "bad news" from ρ'. As a result, the monopolist produces less when she observes "bad news" from ρ'' than when she observes "bad news" from ρ'. In either case, the monopolist makes more extreme decisions when the quality of her information increases.

In Figure 1(a), the rotation of the solid line, $q^M(\cdot;\rho')$, to the dashed line, $q^M(\cdot;\rho'')$, captures the more extreme production decision due to an increase in the quality of information. This in turn induces a mean-preserving spread in the distribution H, as shown by the density function h, in Figure 1(b). Thus, as the monopolist's signal becomes more informative, the distribution of quantities becomes more dispersed while the average quantity produced remains unaffected. Notice, however, that the result makes heavy use of the monopolist's quadratic profit function and the "truth-or-noise" signal, both of which make the problem tractable.

In this paper, we characterize how an increase in the quality of information affects the distribution of optimal actions in a general model with supermodular payoffs. We present (i) an order over the distributions of optimal actions that captures changes in the mean and dispersion, (ii) an order over information structures that captures quality, and (iii) conditions on payoff functions that lead to an equivalence between the two orders.

Consider a third-party who holds preferences over the decision-maker's actions, and has to choose between two information structures, ρ' and ρ''.[2] We say the agent is more *responsive with a higher mean* under ρ'' than under ρ' if any risk-loving third-party prefers the distribution of optimal actions induced by ρ''. Alternatively, we say the agent is more *responsive with a lower mean* under ρ'' than under ρ' if any risk-averse third-party prefers the distribution of optimal actions induced by ρ'. Loosely, responsiveness with a higher mean corresponds to higher variability and higher actions on average (increasing convex stochastic order) while responsiveness with lower mean corresponds to higher variability but lower actions on average (second-order stochastic dominance).

To compare the quality of information structures, we first restrict attention to a class of structures in which higher signal realizations lead to first-order stochastic shifts in posterior beliefs. The restriction is weaker than the common assumption that signals are ordered by the monotone likelihood ratio property (MLRP). Within this restricted class of experiments, we then use the *monotone information order* (Athey and Levin, 2017) to capture quality.[3] Intuitively, information structure ρ'' dominates information structure ρ' in the monotone information order if the signals from ρ'' are more correlated with the state than are the signals from ρ'.

Our main result shows that if ρ'' dominates ρ' in the monotone information order, then an agent, whose payoff function exhibits a supermodular and convex (in actions) marginal utility, is more responsive with a higher mean under ρ'' than under ρ'. Furthermore, we show that if an agent is more responsive with a higher mean under ρ'' than under ρ' for all payoff functions that exhibit supermodular and convex marginal utilities, then ρ'' necessarily dominates ρ' in the monotone information order.

Intuitively, a supermodular payoff function implies that the agent "benefits from matching" her actions to the state, i.e, taking higher actions for higher states of the world and lower actions for lower states. When a payoff function additionally exhibits a supermodular and

[2] In the monopoly example, the third-party could be a social planner who has preferences over the quantity produced by the monopolist. More broadly, we can think of the third-party and the decision-maker as the "sender" and "receiver" in a Bayesian persuasion framework à la Kamenica and Gentzkow (2011).

[3] The monotone information order is equivalent to the supermodular stochastic ordering and the positive dependence ordering when the state is one-dimensional.

convex marginal utility, the agent's "benefit from matching" is non-diminishing as her action increases.[4] A consequence is that the agent's optimal action, as a function of her posterior beliefs, is convex. When the agent's quality of information increases, the distribution of her posterior beliefs becomes more dispersed, which in conjunction with a convex optimal action, leads to a more dispersed distribution of actions with a higher mean.

We also present symmetric results linking responsiveness with a lower mean to payoff functions with a submodular and concave marginal utility. Furthermore, we provide an example in which a higher quality of information does not lead to a more dispersed distribution of actions when the agent's payoff function violates the conditions on the marginal utility. As an application, we reconsider the monopolist example in a more general setting and study how a social planner can regulate the quality of the monopolist's information in order to improve social welfare.

We then extend our comparative statics results to Bayesian games with strategic complementarities. We consider a setting in which different players receive private signals of varying quality about the underlying state of the world before playing a game. Similar to the single agent case, under conditions on the players' marginal utilities, we show that a higher signal quality for any one player leads to a more dispersed distribution of Bayesian Nash equilibrium actions along with an increase or decrease in the mean equilibrium actions for all players.

Our comparative statics results point out a more intricate interaction between a player's equilibrium strategy and the quality of information than has been previously studied. First, we generalize the observation in linear-quadratic games that a player's best-response becomes more dispersed when that player's own signal becomes more informative. Second, when one player's signal becomes more informative about the state, it also becomes (weakly) more correlated (unconditional on the state) to other players' signals. Due to strategic complementarities, a higher quality of information about each others' signal realizations, and thereby actions, implies a more dispersed distribution of best-responses. Third, a player's ex-post desire to match the actions of other players implies that the player has ex-ante incentives to match the distribution of actions of other players. Hence, a player's best-response becomes more dispersed if another player's distribution of actions becomes more dispersed. Our main result shows that the culmination of these three effects is that players are not only responsive to changes in the quality of their own signals but also to changes in the quality of their opponent's signals.

[4]This does not imply that the agent has a non-diminishing marginal utility. Instead, it implies that the marginal gains from matching actions to states dominates the rate at which the marginal utility diminishes. See the discussion following Proposition 1.

The results are fruitful for studying informational externalities in Bayesian games, such as differentiated Bertrand competition and coordination games, under general payoff functions and information structures. For instance, we show that full information sharing between partners in a joint project is optimal without assuming a quadratic payoff function or Gaussian signals.

As an application of our methodology, we then study endogenous information acquisition in Bayesian games with two players. The game is composed of two stages: only player 1 acquires information in the first stage followed by a second stage in which both players choose actions simultaneously. Whether or not player 2 observes player 1's choice of information corresponds to overt and covert information acquisition games respectively. We define the *value of transparency* as the difference in the ex-ante payoffs to player 1 between the overt and the covert games and we show how responsiveness is useful to characterize it. Specifically, we show that the value of transparency is positive or negative depending on (i) the responsiveness of player 2 to player 1's information quality, and (ii) the sign of the externality on player 1 imposed by player 2's action. This in turn has implications on how much information player 1 acquires in the two games.

1.1 Related Literature

Two papers closely related to ours are Jensen (2018) and Lu (2016). Jensen considers a decision-maker who has complete information about the state of the world. His paper characterizes how changes in the distribution over the state of the world affect the induced distribution over optimal actions.[5] In our setting, the decision-maker does not observe the state and the prior distribution over the state is held fixed. Instead, we characterize how changes to the information structure affect the distributions over the posterior beliefs which, in turn, affect the distribution over optimal actions. Lu studies how information acquisition affects choice from a menu. In particular, he shows that a decision-maker has a more dispersed willingness-to-pay for any given menu if the quality of information increases. We instead show that the choice from within a menu becomes more dispersed as the quality of information increases.[6]

This paper also contributes to the literature on monotone comparative statics and the value of information. In the monotone comparative statics literature, our paper is closest to Athey (2002) who characterizes when optimal actions increase as a function of beliefs. We

[5]In the context of our motivating example, the monopolist observes the state θ and optimally produces quantity $q^M(\theta)$. Jensen's paper characterizes how different distributions over θ affect the distribution of $q^M(\theta)$.

[6]Note that there cannot be any meaningful dispersion in choice of action from within a singleton menu. However, the willingness-to-pay for the singleton menu can vary depending on the decision-maker's belief.

take the next step and show how the distribution of optimal actions changes as a function of the distribution over beliefs.[7] Our work also relates to literature on the value of information: Blackwell (1951, 1953), Lehmann (1988), Persico (2000), Quah and Strulovici (2009), and Athey and Levin (2017). In particular, Athey and Levin show that in the class of payoff functions that exhibit complementarities between actions and states, an agent values more information if, and only if, information quality is increasing in the monotone information order. Our results differ from theirs in that we show in the subclass of payoff functions that exhibit supermodular and convex/submodular and concave marginal utilities, the agent's optimal actions are more dispersed if, and only if, information quality is increasing in the monotone information order.

Our comparative statics results for Bayesian games with strategic complementarities are also related to monotone comparative statics of equilibrium actions studied by Vives (1990), Milgrom and Roberts (1994), Villas-Boas (1997), Van Zandt and Vives (2007), as well as the value of information in Bayesian supermodular games studied by Amir and Lazzati (2016). Amir and Lazzati is particularly noteworthy as they show that in supermodular games, a player values more information if information quality is increasing in the monotone information order. Similar to the single agent case, our results differ in that our comparative statics focuses on the distribution of equilibrium actions. We show that in a subclass of supermodular games, the equilibrium actions for all players become more dispersed if information quality for any one player increases in the monotone information order.

Finally, our analysis of the value of *transparency* in Bayesian games is related to the characterization of strategic investment in sequential versus simultaneous games of complete information in Fudenberg and Tirole (1984) and Bulow, Geanakoplos, and Klemperer (1985). We defer a detailed discussion of the relationship to Section 4.

The remainder of the paper is structured as follows: In section 2, we present the single agent framework, and provide sufficient and necessary conditions for an agent to become more responsive as information quality increases. We extend the analysis to Bayesian games with strategic complementarities in Section 3. In Section 4, we present an application to overt and covert information acquisition games and analyze the value of *transparency* in Bayesian games. Section 5 concludes. Proofs that are not presented in the text are in the Appendix.

[7] In the context of our motivating example, Athey (2002) provides comparative statics results on $q^M(s;\rho)$ as a function of the signal realization s for a fixed ρ. We instead provide comparative statics results for the entire mapping $q^M(\cdot;\rho)$ as a function of ρ.

1.2 Preliminary Definitions and Notation

Let X_i, $i = 1, 2, \ldots, m$, and Y be compact subsets of $\mathbb{R}$. Let $X \triangleq \times_{i=1}^m X_i$ be the Cartesian product endowed with the product order so that for $x', x \in X$, $x' \geq x$ if, and only if, $x'_i \geq x_i$ for $i = 1, 2, \ldots, m$. Let $x' \vee x$ denote the join of x' and x, the component-wise maximum, and let $x' \wedge x$ denote the meet of x' and x, the component-wise minimum.

A function $g : X \to \mathbb{R}$ is supermodular (submodular) if $g(x' \vee x) + g(x' \wedge x) \geq (\leq) g(x') + g(x)$ for all $x, x' \in X$. We say that g is modular if it is both supermodular and submodular. We use the terms 'increasing', and 'decreasing' in the weak sense, for example, we say a function $f : Y \to \mathbb{R}$ is increasing if $y' > y$ implies $f(y') \geq f(y)$. We will be explicit when we refer to strict monotonicity. A function $h : X \times Y \to \mathbb{R}$ has increasing (decreasing) differences in $(x; y)$ if for $x' \geq x$, $h(x', y) - h(x, y)$ is increasing (decreasing) in y.

For a differentiable function, $g : X \to \mathbb{R}$, we write $g_{x_i}(x)$ as a shorthand for $\frac{\partial}{\partial x_i} g(x)$ and $g_{x_i x_j}(x)$ for $\frac{\partial^2}{\partial x_i x_j} g(x)$. If g is differentiable and supermodular, then $g_{x_i x_j} \geq 0$ for all $i \neq j$.

2 Single-agent Model

Let $A \triangleq [\underline{a}, \bar{a}]$ be the action space and let $\Theta \triangleq [\underline{\theta}, \bar{\theta}]$ represents the state space. Let $\Delta(\Theta)$ denote the set of all Borel probability measures on Θ. An agent (she) has to choose an action $a \in A$ before observing the realized state of the world $\theta \in \Theta$. The agent's prior belief is denoted by the measure $\mu^o \in \Delta(\Theta)$. We allow for beliefs to be discrete measures with a finite support in Θ or absolutely continuous measures on Θ. Payoffs are given by the function $u : \Theta \times A \to \mathbb{R}$ such that

(A.1) $u(\theta, a)$ is uniformly bounded, measurable in θ, and twice differentiable in a,

(A.2) for all $\theta \in \Theta$, $u(\theta, \cdot)$ is strictly concave in a with $u_{aa}(\theta, \cdot) < 0$,

(A.3) for all $\theta \in \Theta$, there exists an action $a \in A$ such that $u_a(\theta, a) = 0$, and

(A.4) $u(\theta, a)$ is supermodular.

Supermodularity implies that the agent prefers a high action when the state is high and a low action when the state is low. Assumptions (A.1)-(A.3) allow us to characterize the optimal actions by their first order conditions. In Section 2.4, we discuss the difficulties that arise when some of these assumptions are violated.

Given any belief $\mu \in \Delta(\Theta)$, define

$$a^*(\mu) = \arg\max_{a \in A} \int_\Theta u(\theta, a)\mu(d\theta).$$

The compactness of A and the continuity of the utility function guarantee that the solution exists and is measurable. Furthermore, Athey (2002) shows that (A.4) implies $a^*(\mu_2) \geq a^*(\mu_1)$ whenever $\mu_2 \succeq_{FOSD} \mu_1$.[8]

Prior to decision-making, the agent can observe an informative signal about the unknown state. Signals are generated by an information structure $\Sigma_\rho \triangleq \langle S, F(\cdot,\cdot;\rho)\rangle$ where $S \subseteq \mathbb{R}$ is a compact interval, $F(\cdot,\cdot;\rho) : \Theta \times S \to [0,1]$ is a joint probability distribution given by $F(\theta, s; \rho)$, and ρ is an index that is useful when comparing multiple signal structures.

For any information structure Σ_ρ, the marginal of $F(\cdot,\cdot;\rho)$ on Θ, denoted by F_Θ, must satisfy $F_\Theta(\theta) = \int_{\underline{\theta}}^{\theta} \mu^o(d\omega)$. We denote the marginal of $F(\cdot,\cdot;\rho)$ on S by $F_S(\rho)$. Without loss of generality, we assume that all information structures have the same marginal on the signal, i.e., $F_S(s;\rho) = F_S(s)$ for all $s \in S$ and any Σ_ρ. Moreover, F_S has a positive bounded density f_S.[9]

2.1 Order over Distributions of Optimal Actions

From an interim perspective of the decision problem, the agent first observes signal realization $s \in S$ from information structure Σ_ρ, updates her beliefs to a posterior $\mu(\cdot|s;\rho) \in \Delta(\Theta)$ via Bayes rule, and then chooses the optimal action $a^*\big(\mu(\cdot|s;\rho)\big)$. Define the measurable function $a(\rho) : S \to A$ given by $a(s;\rho) = a^*\big(\mu(\cdot|s;\rho)\big)$.

From an ex-ante perspective, the signal realizations are yet to be observed. Therefore, the optimal actions induced by an information structure Σ_ρ are random variables. In particular, $a(\rho)$ is a random variable that is distributed according to $H(\cdot;\rho)$ defined as

$$H(z;\rho) \triangleq F_S\big(\{s : a(s;\rho) \leq z\}\big)$$

[8] For any two beliefs $\mu_1, \mu_2 \in \Delta(\Theta)$, we say that μ_2 first-order stochastically dominates μ_1, denoted $\mu_2 \succeq_{FOSD} \mu_1$, if for any increasing function $g : \Theta \to \mathbb{R}$, $\int_\Theta g(\theta)\mu_2(d\theta) \geq \int_\Theta g(\theta)\mu_1(d\theta)$.

[9] The assumption is without loss of generality: we can apply the integral probability transform to any signal with a continuous marginal distribution $F_S(\rho)$ and create a new signal which is uniformly distributed on the unit interval. The transformed signal still conveys the same information as the original signal. If $F_S(\rho)$ is discontinuous, then, as noted by Lehmann (1988), we can construct a new equally informative signal with a continuous marginal by appropriately distributing the mass at discontinuity points.

for $z \in \mathbb{R}$. The quantile function is defined as

$$\hat{a}(q;\rho) = \inf\{z : q \leq H(z;\rho)\}$$

for $q \in (0,1)$.

Our goal is to characterize how information quality affects the distribution of optimal actions. Thus, the first step is to identify an order over distributions of optimal actions that appropriately captures changes in the mean and dispersion of actions.

Given two information structures $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$, we say that $a(\rho'')$ dominates $a(\rho')$ in the *decreasing convex order* if, for any decreasing convex function $\phi : \mathbb{R} \to \mathbb{R}$

$$\int_{-\infty}^{\infty} \phi(z)dH(z;\rho'') \geq \int_{-\infty}^{\infty} \phi(z)dH(z;\rho').$$

In other words, $a(\rho'')$ dominates $a(\rho')$ in the decreasing convex order if any risk-averse third-party with utility function $-\phi(\cdot)$ prefers $a(\rho')$.

Alternatively, we say that $a(\rho'')$ dominates $a(\rho')$ in the *increasing convex order* if, for any increasing convex function $\varphi : \mathbb{R} \to \mathbb{R}$

$$\int_{-\infty}^{\infty} \varphi(z)dH(z;\rho'') \geq \int_{-\infty}^{\infty} \varphi(z)dH(z;\rho').$$

In other words, $a(\rho'')$ dominates $a(\rho')$ in the increasing convex order if any risk-loving third-party with utility function $\varphi(\cdot)$ prefers $a(\rho'')$. Note that if $a(\rho'')$ dominates $a(\rho')$ in both the decreasing convex and increasing convex order, then $a(\rho'')$ is a mean-preserving spread of $a(\rho')$.

Responsiveness: Given two information structures $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$, we say that

i. an agent is **more responsive with a lower mean** under $\Sigma_{\rho''}$ than under $\Sigma_{\rho'}$ if, and only if, $a(\rho'')$ dominates $a(\rho')$ in the **decreasing convex order**, and

ii. an agent is **more responsive with a higher mean** under $\Sigma_{\rho''}$ than under $\Sigma_{\rho'}$ if, and only if, $a(\rho'')$ dominates $a(\rho')$ in the **increasing convex order**.

In other words, the agent is more responsive under information structure $\Sigma_{\rho''}$ than under $\Sigma_{\rho'}$ if the information she receives from $\Sigma_{\rho''}$ leads the agent to take more varied actions. The definition of responsiveness connects an agent's behavior under different informational sources to changes in the dispersion and expectation of the agent's optimal action. As a short hand, we say the agent is responsive if she is either more responsive with a higher mean or more responsive with a lower mean.

Lemma 1 below provides equivalent characterizations of responsiveness. The first equivalence provides an alternate definition by comparing the distribution functions of the optimal actions. The second equivalence characterizes responsiveness as a comparison of the quantile functions. These alternative definitions are particularly useful when the optimal actions are monotone in the signal realization, a natural consequence when payoffs are supermodular and beliefs are ordered by first-order stochastic dominance.

Lemma 1 (Shaked and Shantikumar, 2007; Theorem 4.A.2-3)
Given two information structures $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$, the following are equivalent:

i. An agent is more responsive with lower mean under $\Sigma_{\rho''}$ than under $\Sigma_{\rho'}$.

ii. For all $x \in \mathbb{R}$,

$$\int_{-\infty}^{x} H(z;\rho')dz \leq \int_{-\infty}^{x} H(z;\rho'')dz.$$

iii. For all $t \in [0,1]$,

$$\int_{0}^{t} \hat{a}(q;\rho')dq \geq \int_{0}^{t} \hat{a}(q;\rho'')dq.$$

Similarly, the following are equivalent:

iv. An agent is more responsive with higher mean under $\Sigma_{\rho''}$ than under $\Sigma_{\rho'}$.

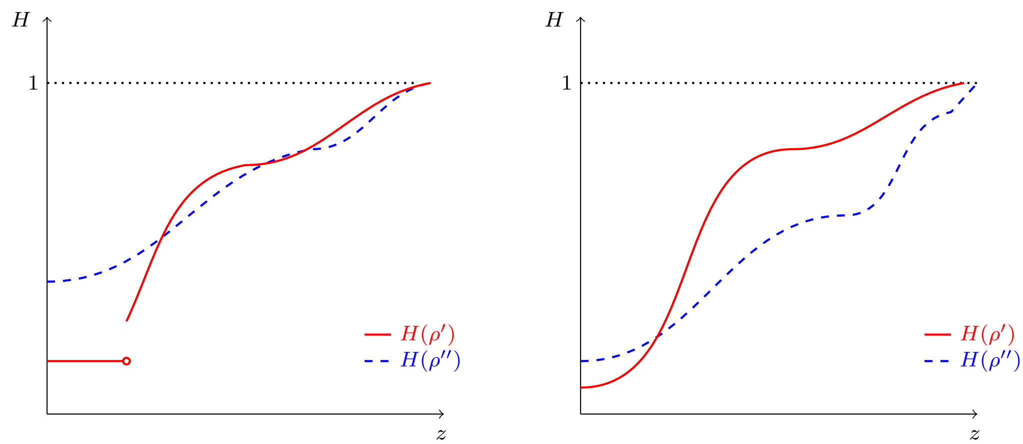
v. For all $x \in \mathbb{R}$,

$$\int_{x}^{\infty} H(z;\rho'')dz \leq \int_{x}^{\infty} H(z;\rho')dz.$$

vi. For all $t \in [0,1]$,

$$\int_{t}^{1} \hat{a}(q;\rho'')dq \geq \int_{t}^{1} \hat{a}(q;\rho')dq.$$

Figure 2 below plots the distribution over actions induced by two information structures $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$. In Figure 2(a), the area between the y-axis and $H(\rho'')$ (the dashed curve) is smaller than that of $H(\rho')$ (the solid curve) which implies the mean of the optimal actions induced by $\Sigma_{\rho''}$ is lower than the mean induced by $\Sigma_{\rho'}$. Furthermore, integrating $H(z;\rho') - H(z;\rho'')$ left to right always yields a negative value which, by Lemma 1.*ii*, implies responsiveness with a lower mean. In contrast, in Figure 2(b), the area between the y-axis and $H(\rho'')$ is bigger than that of $H(\rho')$ which implies the mean of the optimal actions induced by $\Sigma_{\rho''}$ is larger than the mean induced by $\Sigma_{\rho'}$. Furthermore, integrating $H(z;\rho') - H(z;\rho'')$ right to left always yields a positive value which, by Lemma 1.v, implies responsiveness with a higher mean.



(a) Responsiveness with a lower mean

(b) Responsiveness with a higher mean

Figure 2: CDF of Optimal Actions and Responsiveness

2.2 The Monotone Information Order

The next step is to determine an appropriate way to compare different information structures. We first restrict attention to information structures in which higher signal realizations lead to a first-order stochastic increase in beliefs. This assumption is weaker than the monotone likelihood ratio property commonly assumed in settings with complementarities.

(A.5) For any given information structure Σ_ρ, $s' > s$ implies $\mu(\cdot|s';\rho) \succeq_{FOSD} \mu(\cdot|s;\rho)$.

Monotone Information Order: $\Sigma_{\rho''}$ dominates $\Sigma_{\rho'}$ in the monotone information order, denoted $\rho'' \succeq_{MIO} \rho'$, if for all $q \in [0,1]$

$$\mu\big(\cdot\,|F_S(s) \geq q;\rho''\big) \succeq_{FOSD} \mu\big(\cdot\,|F_S(s) \geq q;\rho'\big)$$

and

$$\mu\big(\cdot\,|F_S(s) \leq q;\rho'\big) \succeq_{FOSD} \mu\big(\cdot\,|F_S(s) \leq q;\rho''\big).$$

Intuitively, when $\rho'' \succeq_{MIO} \rho'$, the signal and the state are more positively correlated under $\Sigma_{\rho''}$ than under $\Sigma_{\rho'}$, and consequently, the agent updates her beliefs more "aggressively" under

$\Sigma_{\rho''}$. More formally, by (A.5), high signal realizations are evidence of high states. The agent considers a signal realization above the q^{th} quantile from $\Sigma_{\rho''}$ as a stronger evidence that the state could be high (than a signal realization above the q^{th} quantile from $\Sigma_{\rho'}$). Consequently, the agent is more optimistic when she observes a signal realization above the q^{th} quantile from $\Sigma_{\rho''}$ than from $\Sigma_{\rho'}$. Similarly, a signal realization below the q^{th} quantile from $\Sigma_{\rho''}$ is a stronger evidence that the state could be low (than a signal realization below the q^{th} quantile from $\Sigma_{\rho'}$). Thus, the agent is more pessimistic when she observes a signal realization below the q^{th} quantile from $\Sigma_{\rho''}$ than from $\Sigma_{\rho'}$.

Example 1: Truth-or-Noise signals
To avoid confusion, let $(\tilde{\theta}, \tilde{s})$ be the random variables representing the state and the signal while (θ, s) represent typical realizations of the random variables. Σ_ρ belongs to a class of information structures such that with probability $\rho \in [0, 1]$, the signal reveals the state ($\tilde{s} = \tilde{\theta}$), and with probability $1 - \rho$, the signal and the state are identically and independently distributed. Thus, with probability $1 - \rho$, the signal is uninformative. Then, $\rho'' \succeq_{MIO} \rho'$ if $\rho'' > \rho'$.

Example 2: Normal prior and signals:
Σ_ρ belongs to a class of information structures such that the state and the signal are multivariate normally distributed random variables with

$$\begin{bmatrix} \tilde{\theta} \\ \tilde{s} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \theta_0 \\ \theta_0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right)$$

where $\rho \in [0, 1]$ is the correlation coefficient between the two normally distributed variables. Hence, given a signal realization s, the agent's posterior is given by

$$\tilde{\theta}|s \sim \mathcal{N}\Big(\rho s + \theta_0(1 - \rho),\ 1 - \rho\Big)$$

When ρ increases, the agent's posterior (i) places more weight on the observed signal as evidenced by the shift in the mean and (ii) reduces the variance in the agent's posterior. Then, $\rho'' \succeq_{MIO} \rho'$ if $\rho'' > \rho'$.

In the Appendix, we discuss why the monotone information order is the relevant order to consider when characterizing responsiveness. We also elaborate how the monotone information order compares to the more familiar Blackwell informativeness (Blackwell, 1951, 1953) or the

Lehmann (accuracy) order (Lehmann, 1988).[10]

2.3 Monotone Information Order and Responsiveness

The main contribution of this paper is to identify a class of decision problems for which the agent becomes more responsive when information quality increases according to the monotone information order. Let $\mathcal{U}^{\uparrow}$ be the class of payoff functions $u : \Theta \times A \to \mathbb{R}$ that satisfy (A.1)-(A.4) and have a marginal utility $u_a(\theta, a)$ that is

i. convex in a for all $\theta \in \Theta$, and

ii. supermodular in (θ, a).

Below, we show that an agent with a payoff function $u \in \mathcal{U}^{\uparrow}$ becomes more responsive with a higher mean (hence the up arrow) as information quality increases in the monotone information order.

Let $\mathcal{U}^{\downarrow}$ be the class of payoff functions $u : \Theta \times A \to \mathbb{R}$ that satisfy (A.1)-(A.4) and have a marginal utility $u_a(\theta, a)$ that is

i. concave in a for all $\theta \in \Theta$, and

ii. submodular in (θ, a).

Below, we show that an agent with a payoff function $u \in \mathcal{U}^{\downarrow}$ becomes more responsive with a lower mean (hence the down arrow) as information quality increases in the monotone information order.

Theorem 1 *Consider two information structures $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$ that satisfy* (A.5)*. $\Sigma_{\rho''}$ dominates $\Sigma_{\rho'}$ in the monotone information order if, and only if, an agent with any payoff $u \in \mathcal{U}^{\uparrow}$ $\left[u \in \mathcal{U}^{\downarrow}\right]$, is more responsive with a higher* [*lower*] *mean under $\Sigma_{\rho''}$ than under $\Sigma_{\rho'}$.*

Intuitively, the higher the quality of the agent's information, the more aggressively she updates her belief. Thus, when information quality increases in the monotone information order, the agent's posterior beliefs becomes more dispersed. Theorem 1 provides the conditions on the agent's utility function under which we can map the more dispersed distribution of posterior beliefs to a more dispersed distribution of actions that incorporates monotone changes

[10] See Persico (2000) and Jewitt (2006) for detailed description and applications. Definitions can be found in the Appendix.

to the mean optimal action. The mechanism behind Theorem 1 is best understood through Proposition 1 below, which shows that a payoff function $u \in \mathcal{U}^{\uparrow}$ $\left[u \in \mathcal{U}^{\downarrow}\right]$ leads to optimal actions that are "convex" ["concave"] in the agent's posterior belief. We then show how this convexity/concavity interacts with the quality of information to result in more dispersed actions.

Proposition 1 *Let $\mu_1, \mu_2 \in \Delta(\Theta)$ be any two beliefs with $\mu_2 \succeq_{FOSD} \mu_1$. If $u \in \mathcal{U}^{\uparrow}$, then for any $\lambda \in [0,1]$*

$$a^*\big(\lambda\mu_1 + (1-\lambda)\mu_2\big) \leq \lambda a^*(\mu_1) + (1-\lambda)a^*(\mu_2)$$

If $u \in \mathcal{U}^{\downarrow}$, the opposite inequality holds.

Proof. Let $a_i = a^*(\mu_i)$ for $i = 1, 2$, $a_\lambda = \lambda a_1 + (1-\lambda)a_2$, and $\mu_\lambda = \lambda\mu_1 + (1-\lambda)\mu_2$. By the first order condition, we have that $\int_\Theta u_a(\theta, a_i)\mu_i(d\theta) = 0$. Let $u \in \mathcal{U}^{\uparrow}$.

$$\begin{aligned}
\int_\Theta u_a(\theta, a_\lambda)\mu_\lambda(d\theta) &\leq \lambda \int_\Theta u_a(\theta, a_1)\mu_\lambda(d\theta) + (1-\lambda)\int_\Theta u_a(\theta, a_2)\mu_\lambda(d\theta) \\
&= \lambda^2 \int_\Theta u_a(\theta, a_1)\mu_1(d\theta) + (1-\lambda)^2 \int_\Theta u_a(\theta, a_2)\mu_2(d\theta) \\
&+ \lambda(1-\lambda)\left[\int_\Theta u_a(\theta, a_2)\mu_1(d\theta) + \int_\Theta u_a(\theta, a_1)\mu_2(d\theta)\right] \\
&= \lambda(1-\lambda)\int_\Theta \left[u_a(\theta, a_1) - u_a(\theta, a_2)\right]\left(\mu_2(d\theta) - \mu_1(d\theta)\right) \\
&\leq 0
\end{aligned}$$

where the first inequality follows from the convexity of u_a. As already noted, supermodularity of the utility $u(\theta, a)$ along with $\mu_2 \succeq_{FOSD} \mu_1$ implies $a_2 \geq a_1$. By supermodularity of the marginal utility u_a, we have $u_a(\theta, a_1) - u_a(\theta, a_2)$ is a decreasing function of θ. The last inequality then follows from the definition of first-order stochastic dominance. Since the marginal value of a_λ is non-positive at μ_λ, we must have $a^*(\mu_\lambda) \leq a_\lambda$. A symmetric argument establishes that if $u \in \mathcal{U}^{\downarrow}$, then $a^*(\mu_\lambda) \geq a_\lambda$. ■

Henceforth, we focus on payoffs in $\mathcal{U}^{\uparrow}$ but the arguments we provide can be symmetrically applied to payoffs in $\mathcal{U}^{\downarrow}$. To see the intuition behind the additional assumptions on the utility functions in $\mathcal{U}^{\uparrow}$, consider a " revision" process by which an agent starts at some arbitrary action $\hat{a} \in A$ and adjusts this action as the state changes. Supermodularity of u implies that the state and the action are complements, i.e., for two states $\theta > \theta'$, the difference in the marginal utility

$u_a(\theta, a) - u_a(\theta', a)$ is non-negative. Therefore, the agent is willing to adjust $\hat{a}$ upwards as the state increases.

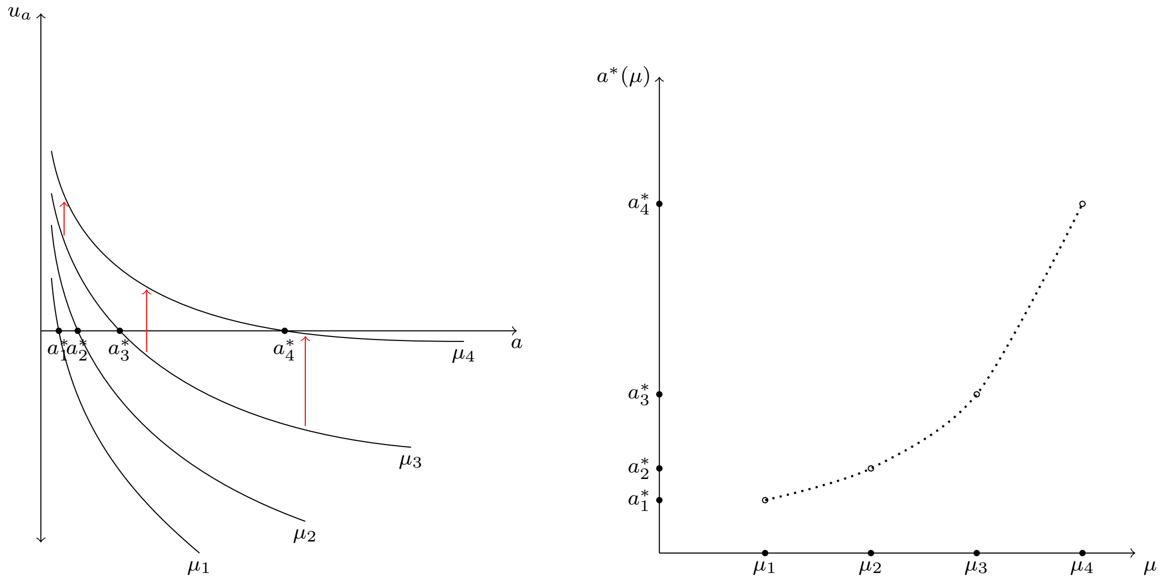
However, supermodularity of u does not tell us anything about the strength of the complementarities between the action and the state. It could be that complementarities are negligible for low actions but substantial for high actions. In such a case, the agent's payoff additionally satisfies *increasing supermodularity*, i.e., $u_a(\theta, a) - u_a(\theta', a)$ is non-negative and increasing in a. Therefore, if $\hat{a}$ is small, $u_a(\theta, \hat{a}) - u_a(\theta', \hat{a})$ is also small and the agent is only willing to adjust $\hat{a}$ upwards by a negligible amount as the state increases. Conversely, if $\hat{a}$ is large, $u_a(\theta, \hat{a}) - u_a(\theta', \hat{a})$ is also large and the agent is willing to adjust $\hat{a}$ upwards by a substantial amount.

On the other hand, the concavity of the agent's payoff function implies that she has a diminishing marginal utility. An agent who starts the "revision" process at a small $\hat{a}$ is more willing to increase her action than an agent who starts at a high $\hat{a}$. Thus, there are two opposing forces at work. When the agent's payoff u belongs to $\mathcal{U}^{\uparrow}$, the rate at which her marginal utility diminishes is less than the rate at which the complementarities between her action and the state increase.

For a simple visual representation, let the state space be $\Theta = \{\underline{\theta}, \bar{\theta}\}$ with $\bar{\theta} > \underline{\theta}$. With some abuse of notation, let $\mu = \mathbb{P}(\theta = \bar{\theta}) \in [0, 1]$ represent the agent's belief that $\theta = \bar{\theta}$. Consider four different beliefs $\{\mu_i\}_{i=1,2,3,4}$ such that, $\mu_{i+1} = \mu_i + \delta$ for some $\delta > 0$. Figure 3(a) below plots out the expected marginal utility of a payoff function $u \in \mathcal{U}^{\uparrow}$ for the different beliefs. Since the payoff is concave in a, the marginal utilities are downward sloping. The optimal action $a^*(\mu_i)$ is given by the action at which the expected marginal utility under belief μ_i intersects the x-axis. Since $\mu_4 > \mu_3 > \mu_2 > \mu_1$, the beliefs are ordered by first-order stochastic dominance with $\mu_4 \succeq_{FOSD} \mu_3 \succeq_{FOSD} \mu_2 \succeq_{FOSD} \mu_1$. Supermodularity implies that the expected marginal utility of μ_i always lies below the expected marginal utility of μ_{i+1}. Thus, $a^*(\mu_4) \geq a^*(\mu_3) \geq a^*(\mu_2) \geq a^*(\mu_1)$.

Furthermore, *increasing supermodularity* implies that the gap between the expected marginal utilities of μ_{i+1} and μ_i is widening as the action increases. We capture this by showing that the height of the red arrows increases left to right. Finally, the marginal utilities themselves are convex curves which implies that the marginal utility diminishes at a diminishing rate. All these properties combined result in $a^*(\mu_4) - a^*(\mu_3) > a^*(\mu_3) - a^*(\mu_2) > a^*(\mu_2) - a^*(\mu_1)$. Figure 3(b) depicts this "convexity" property as described in Proposition 1.

To see how the "convexity" of the optimal action is related to responsiveness, let us continue with the above simplified setting with two states. Let $\mu_o \in (0, 1)$ be the agent's prior belief



(a) Marginal utilities (b) Optimal action

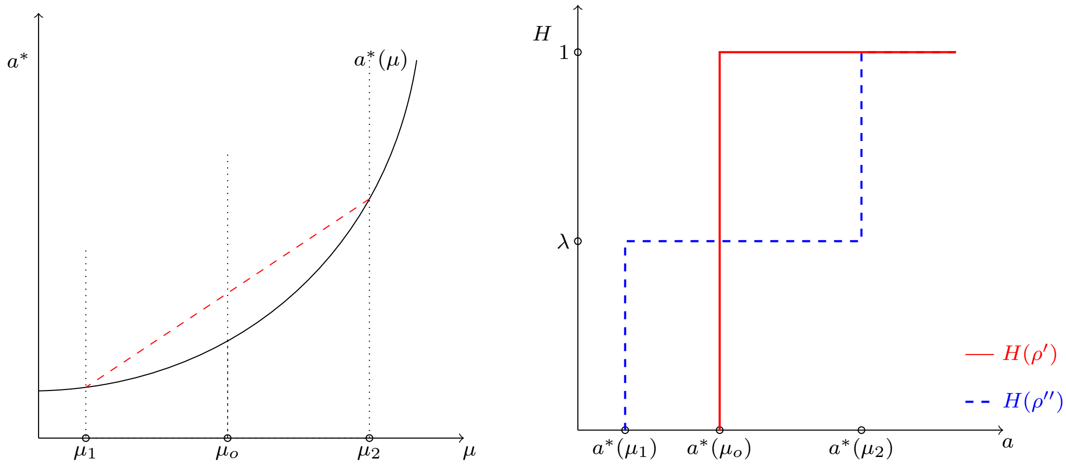
Figure 3: Convexity for $u \in \mathcal{U}^{\uparrow}$

that the state is $\bar{\theta}$ and let $\Sigma_{\rho'}$ be a completely uninformative information structure. Then, $\Sigma_{\rho'}$ induces $a^*(\mu_o)$ with probability one.

Let $\Sigma_{\rho''}$ be an information structure that induces two posteriors $\{\mu_1, \mu_2\}$ with probability $\{\lambda, 1-\lambda\}$. Without loss of generality, assume $\mu_2 > \mu_1$ which implies $\mu_2 \succeq_{FOSD} \mu_1$. Consistency of Bayes-updating implies $\mu_o = \lambda\mu_1 + (1-\lambda)\mu_2$. $\Sigma_{\rho''}$ induces optimal actions $a^*(\mu_1)$ with probability λ and $a^*(\mu_2)$ with probability $1-\lambda$. Furthermore, given the supermodularity of $u(\theta, a)$ and $\mu_2 \succeq_{FOSD} \mu_1$, $a^*(\mu_2) \geq a^*(\mu_1)$.

From Proposition 1, if $u \in \mathcal{U}^{\uparrow}$, then $\lambda a^*(\mu_1) + (1-\lambda)a^*(\mu_2) \geq a^*(\lambda\mu_1 + (1-\lambda)\mu_2) = a^*(\mu_o)$. In Figure 4(a) below, the average action from the more informative structure $\Sigma_{\rho''}$ is given by the point on the dashed line directly above μ_o while the average action from the uninformative structure $\Sigma_{\rho'}$ is given by the point on the solid curve directly above μ_o.

Figure 4(b) maps the induced distribution over optimal actions. The dashed line, $H(\rho'')$, maps the distribution of actions under $\Sigma_{\rho''}$ with a mass of size λ at $a^*(\mu_1)$ and another mass of size $1-\lambda$ at $a^*(\mu_2)$. Similarly, the solid line, $H(\rho')$, maps the distribution of actions under $\Sigma_{\rho'}$ which places all the mass at $a^*(\mu_o)$. Notice the integral $\int_x^{\infty} H(z;\rho'') - H(z;\rho')dz \leq 0$ for all $x \in \mathbb{R}$ which implies, by Lemma 1.v, that the agent is more responsive with a higher mean under $\Sigma_{\rho''}$ than under $\Sigma_{\rho'}$.



(a) Optimal action (b) Induced distribution

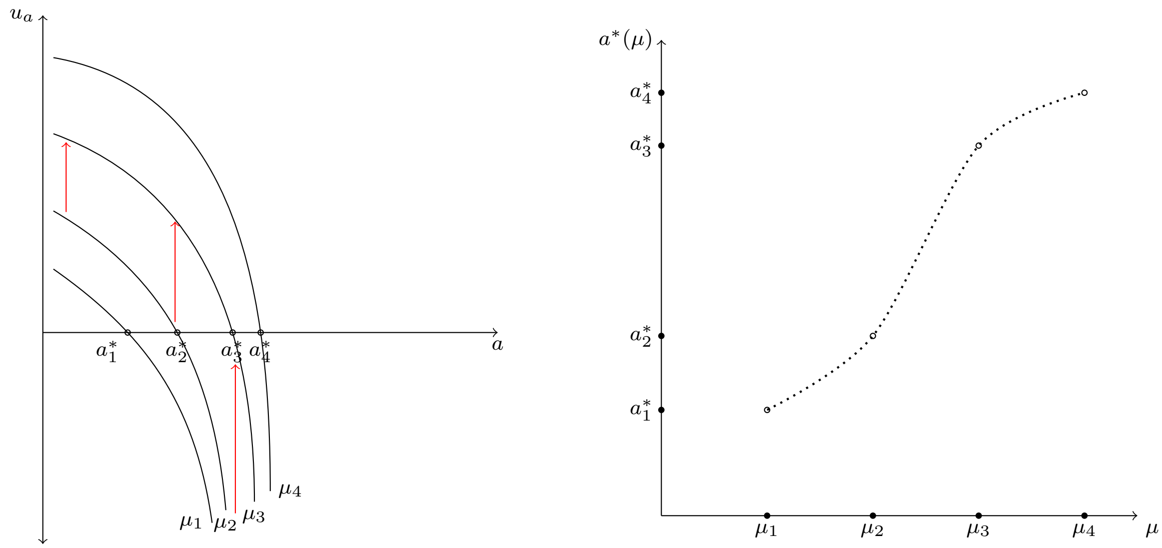
Figure 4: Convexity of a^* and responsiveness with higher mean

2.4 Non-responsive Optimal Actions

In this section, we explore why a higher quality of information may not lead to more responsive optimal actions when $u \notin \mathcal{U}^{\uparrow} \cup \mathcal{U}^{\downarrow}$. Once again, let the state space be $\Theta = \{\underline{\theta}, \bar{\theta}\}$. Consider four different beliefs $\{\mu_i\}_{i=1,2,3,4}$ such that $\mu_{i+1} = \mu_i + \delta$ for some $\delta > 0$. Once again, beliefs are ordered by first-order stochastic dominance with $\mu_4 \succeq_{FOSD} \mu_3 \succeq_{FOSD} \mu_2 \succeq_{FOSD} \mu_1$.

Figure 5 below shows why *increasing supermodularity* alone is not sufficient to get the convexity property from Proposition 1. In Figure 5(a), we plot the expected marginal utilities of some payoff function u. Notice that supermodularity still holds $-$ the expected marginal utility of μ_i lies below the expected marginal utility of μ_{i+1}. Thus, $a^*(\mu_{i+1}) \geq a^*(\mu_i)$. Furthermore, *increasing supermodularity* still holds $-$ the height of the red arrows increases left to right. However, the marginal utilities are now concave which implies that the marginal utility diminishes at an accelerating rate. Hence, $u \notin \mathcal{U}^{\uparrow}$. Furthermore, $a^*(\mu_4) - a^*(\mu_3) < a^*(\mu_3) - a^*(\mu_2)$ whereas $a^*(\mu_3) - a^*(\mu_2) > a^*(\mu_2) - a^*(\mu_1)$. Figure 5(b) depicts this "non-convexity" of the optimal action as a function of beliefs.

Figure 6 below illustrates why the agent may not be responsive to an increase in the quality of information when the optimal action is neither convex nor concave, as in Figure 5(b). Let $\Sigma_{\rho''}$ be an information structure that induces three posteriors $\{\mu_1, \mu_o, \mu_2\}$ with equal probability such that $\mu_2 \succeq_{FOSD} \mu_o \succeq_{FOSD} \mu_1$. Let $\Sigma_{\rho'}$ induce posteriors $\{\mu_1, \mu_2, \mu_3, \mu_4\}$ with probability $\{1/6, 1/6, 1/3, 1/3\}$ such that $\mu_4 \succeq_{FOSD} \mu_3 \succeq_{FOSD} \mu_2 \succeq_{FOSD} \mu_1$. Notice that $\Sigma_{\rho'}$ is a garbling



(a) Marginal utilities

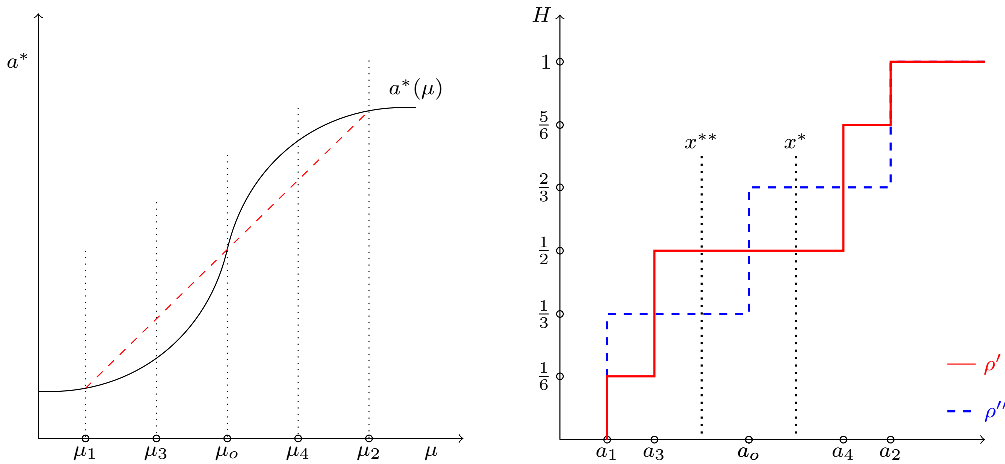
(b) Optimal action

Figure 5: Non-convexity for $u \notin \mathcal{U}^{\uparrow}$

of $\Sigma_{\rho''}$ and thus, $\rho'' \succeq_{MIO} \rho'$.[11]

Let $a^*(\mu)$ be neither convex nor concave and let the average action under $\Sigma_{\rho''}$ equal the average action under $\Sigma_{\rho'}$. In Figure 6(a) below, this corresponds to the point of intersection of the dashed line and the solid curved line at μ_o. Figure 6(b) maps the distribution over optimal actions. $\Sigma_{\rho''}$ induces the dashed line while $\Sigma_{\rho'}$ induces the solid line. If we start integrating from the right, then $\int_x^\infty H(z;\rho'') - H(z;\rho')dz \leq 0$ for all $x > a^*(\mu_4)$ but the sign changes at some point $x^* \in (a^*(\mu_o), a^*(\mu_4))$. If we integrate from the left, then $\int_{-\infty}^x H(z;\rho'') - H(z;\rho')dz \geq 0$ for all $x < a^*(\mu_3)$ but the sign changes at some point $x^{**} \in (a^*(\mu_3), a(\mu_o))$.

[11] A garbling is a kernel $Q : S \times S \to [0,1]$ so that $F(s'|\theta;\rho') = \int_{s\in S} Q(s'|s)dF(s|\theta;\rho'')$



(a) Optimal action (b) Induced distribution

Figure 6: Non-convexity/concavity and non-responsiveness

We can therefore conclude that the agent is neither more responsive with a higher mean nor more responsive with a lower mean. In fact, as the average action under $\Sigma_{\rho''}$ equals the average action under $\Sigma_{\rho'}$, we can conclude that $a(\rho'')$ and $a(\rho')$ cannot be ordered by most univariate stochastic variability orders such as second-order stochastic dominance, mean-preserving spreads, Lorenz order, dilation order, and dispersive order.[12]

Another reason why a higher quality of information may not lead to more responsiveness is when the interior solution assumption, (A.3), is violated. Suppose the upper limit on the action space, $\bar{a}$, is a binding constraint for the prior, i.e., $a^*(\mu_o) = \bar{a}$. Let $\Sigma_{\rho'}$ be a completely uninformative information structure. Then, $\Sigma_{\rho'}$ induces $\bar{a}$ with probability one, thereby first-order stochastically dominating the distribution over actions induced by any other information structure $\Sigma_{\rho''}$, even if $\rho'' \succeq_{MIO} \rho'$.

2.5 Application: Pigouvian Subsidies and Monopoly Production

In the Introduction, we considered the effect of information quality on a monopolist's production decision in a highly stylized example. In this section, we consider the example in a more general setting as follows: a monopolist who produces $q \in [0, \bar{q}]$ faces a downward sloping inverse demand curve $P(q)$ and a cost function $c(\theta, q)$ where the parameter $\theta \in \Theta$ is unknown. The monopolist holds a prior $\mu^o \in \Delta(\Theta)$. As θ increases, the marginal cost declines, i.e. $c(\theta, q)$ is submodular in (θ, q). We assume that the monopolist's profit $\pi(\theta, q) = qP(q) - c(\theta, q)$ is

[12]Shaked and Shanthikumar (2007) provide a thorough treatment of these orders.

strictly concave in q and admits an interior solution for each θ.

Prior to making any production decisions, the monopolist can acquire an information structure from a set of experiments $\{\Sigma_\rho\}_{\rho\in\mathbb{R}}$ at cost $\kappa(\rho)$. We assume the experiments are totally ordered by the monotone information order so that $\rho'' > \rho'$ implies $\rho'' \succeq_{MIO} \rho'$ and $\kappa(\rho'') \geq \kappa(\rho')$.

Consider a social planner who is unable to regulate prices or quantities but can influence the quality of information the monopolist acquires. For example, the social planner may subsidize the monopolist's cost of information acquisition or place a cap on the quality of information that can be acquired. Should the social planner encourage of discourage information acquisition by the monopolist?[13]

Given a choice of information structure Σ_ρ and a signal realization $s \in S$, the monopolist updates her belief to the posterior $\mu(\cdot|s;\rho)$ and produces the monopolist optimal quantity $q^M(s;\rho)$. Thus, the monopolist's ex-ante problem is to choose an information structure that maximizes

$$\int_S \int_\Theta \pi\big(\theta, q^M(s;\rho)\big)\mu(d\theta|s;\rho)dF_S(s) - \kappa(\rho).$$

In contrast, the social planner takes the consumer surplus into account. Let $CS(q)$ be the consumer surplus when the monopolist produces q. The planner's ex-ante payoff given an information structure Σ_ρ is given by

$$\int_S \int_\Theta \pi\big(\theta, q^M(s;\rho)\big)\mu(d\theta|s;\rho)dF_S(s) + \int_S CS\big(q^M(s;\rho)\big)dF_S(s) - \kappa(\rho).$$

Thus, the planner has a higher demand for information than the monopolist if a higher quality of information increases the expected consumer surplus, i.e., information is a positive externality on the consumer even if the unknown parameter θ has no direct effect on consumer welfare.

Proposition 2 *Let $-qP''(q)/P'(q) \leq 1$ and let the profit function $\pi \in \mathcal{U}^\uparrow$. Then the social planner has a* higher *demand for information than the monopolist.*

Intuitively, the assumption that $-qP''(q)/P'(q) \leq 1$ implies that as production increases, the consumers capture more and more of the welfare gains than does the monopolist. Therefore, the consumer surplus is a convex function of the quantity produced, which in turn implies that consumers benefit as the monopolist's production becomes more responsive with higher mean. From Theorem 1, we get the desired responsiveness behavior when $\pi \in \mathcal{U}^\uparrow$.

[13] Athey and Levin (2017) consider a similar porblem. However, in their application, the planner can regulate prices/quantities as well as the quality of information.

3 Supermodular Games

In this section, we extend our results from the single-agent framework to supermodular games with incomplete information. This class of games includes beauty contests, oligopolistic competition, games with network effects, search models, and investment games. It is useful to understand how information quality affects the equilibrium of these games in a general setting.

3.1 Setup

There are n players with $N \triangleq \{1, 2, \dots, n\}$ denoting the set of players. Let $\Theta_i \triangleq [\underline{\theta}_i, \bar{\theta}_i]$ be the state space for player i. Let $\Theta = \times_{i \in N} \Theta_i$ and $\Theta_{-i} = \times_{j \neq i} \Theta_j$. The players hold a common prior $\mu^o \in \Delta(\Theta)$. Let $F_{\Theta_i} : \Theta_i \to [0, 1]$ be the marginal on Θ_i and $F_{\Theta_{-i}}(\cdot|\theta_i) : \Theta_{-i} \to [0, 1]$ be the joint distribution on Θ_{-i} conditional on state $\theta_i \in \Theta_i$ induced by μ^o. Once again, we allow for beliefs to be discrete measures with finite support in Θ, absolutely continuous, or a mixture. Additionally, we assume that

(A.6) for all $i \in N$, $\theta_i' > \theta_i$ implies $F_{\Theta_{-i}}(\cdot|\theta_i') \succeq_{FOSD} F_{\Theta_{-i}}(\cdot|\theta_i)$

which is a weaker assumption than affiliation. Notice that our setup accommodates games of independent or common values.

Let $A_i \triangleq [\underline{a}_i, \bar{a}_i]$ be the action space of player i. Let $A = \times_{i \in N} A_i$ and $A_{-i} = \times_{j \neq i} A_j$. The payoff for each player $i = 1, ..., n$ is given by a utility function $u^i : \Theta_i \times A \to \mathbb{R}$ such that

(A.7) $u^i(\theta_i, a)$ is uniformly bounded, measurable in θ_i, continuous in a, and twice differentiable in a_i,

(A.8) for all $(\theta_i, a_{-i}) \in \Theta_i \times A_{-i}$, $u^i(\theta_i, a_{-i}, \cdot)$ is strictly concave in a_i,

(A.9) for all $(\theta_i, a_{-i}) \in \Theta_i \times A_{-i}$, there exists an action $a_i \in A_i$ such that $u^i_{a_i}(\theta_i, a_{-i}, a_i) = 0$, and

(A.10) $u^i(\theta, a)$ has increasing differences in $(\theta_i, a_{-i}; a_i)$.

Similar to the single-agent framework, (A.10) implies that there are complementarities between the state of the world and a player's action. Additionally, there are now strategic complementarities between the players' actions. Thus, when player j takes a higher action, player i wants to do the same.

Following the terminology introduced by Bergemann and Morris (2016), we decompose the entire game of incomplete information into two components: the basic game and the information structure. The basic game $G \triangleq (N, \{A_i, u^i\}_{i\in N}, \mu^o)$ is composed of (i) a set of players N, (ii) for each player $i \in N$, an action space A_i along with a payoff function $u^i : \Theta_i \times A \to \mathbb{R}$, and (iii) a common prior $\mu^o \in \Delta(\Theta)$.

The second component of the Bayesian game is the information structure $\Sigma_\rho = \times_{i\in N}\Sigma_{\rho_i}$ where for each player $i = 1, ..., n$, signals are generated by $\Sigma_{\rho_i} \triangleq \big(S_i, F(\rho_i)\big)$. $S_i \subseteq \mathbb{R}$ is a compact signal space, $F(\rho_i) : \Theta_i \times S_i \to [0, 1]$ is a joint probability distribution over $\Theta_i \times S_i$ given by $F(\theta_i, s_i; \rho_i)$, and ρ_i is an index.[14] Let $F_{S_i}(\rho_i) : S_i \to [0, 1]$ be the marginal on S_i. Once again, we assume without loss of generality that for any information structure Σ_{ρ_i}, $F_{S_i}(s_i; \rho_i) = F_{S_i}(s_i)$ for all $s_i \in S_i$. Moreover, F_{S_i} has a positive and bounded density f_{S_i}.

Let $S = \times_{i\in N}S_i$. An information structure Σ_ρ induces a joint distribution over $\Theta \times S$ which we denote by $\boldsymbol{F}(\theta, s; \rho)$. The following are working assumptions for this section:

(A.11) For all $s \in S$ and $\theta \in \Theta$, $\boldsymbol{F}(s|\theta; \rho) = \prod_{i\in N} F(s_i|\theta_i; \rho_i)$.

(A.12) For all players $i \in N$, $s'_i > s_i$ implies $\mu(\cdot|s'_i; \rho_i) \succeq_{FOSD} \mu(\cdot|s_i; \rho_i)$.

(A.13) For all players $i \in N$, $\theta'_i > \theta_i$ implies $F(\cdot|\theta'_i; \rho_i) \succeq_{FOSD} F(\cdot|\theta_i; \rho_i)$.

Assumption (A.11) implies that player i can directly learn about θ_i but cannot directly learn about other players' states, θ_{-i}, or signal realizations, s_{-i}. Assumption (A.12) is an extension of (A.5) and implies that higher signal realizations lead to a first-order increase in a player's belief. Assumption (A.13) implies the converse: higher states are likely to lead to higher signal realizations. A distribution over the state and signal space that satisfies the monotone likelihood ratio property also satisfies (A.12)-(A.13).

The full game of incomplete information is given by $\mathcal{G}_\rho \triangleq (\Sigma_\rho, G)$. Both components of the game are common knowledge. First, each player $i \in N$ privately observes a signal realization $s_i \in S_i$ generated from Σ_{ρ_i} and updates her belief to $\mu(\cdot|s_i; \rho_i) \in \Delta(\Theta)$. Then, the players participate in the basic game G by simultaneously choosing an action.

Momentarily ignoring existence issues, let $a^\star(\rho) = \big(a^\star_1(\rho), a^\star_2(\rho), \ldots, a^\star_n(\rho)\big)$ be a profile of pure strategy actions that constitute a Bayesian Nash equilibrium (BNE) of the game $\mathcal{G}_\rho$, and let $a^\star_{-i}(\rho)$ be the profile of BNE strategies excluding player i. For each player $i \in N$, $a^\star_i(\rho) : S_i \to A_i$ is a measurable function. We interpret $a^\star_i(s_i; \rho)$ as the solution to

$$\max_{a_i\in A_i} \int_{\Theta\times S_{-i}} u^i\big(\theta_i, a^\star_{-i}(s_{-i}; \rho), a_i\big) d\boldsymbol{F}(\theta, s_{-i}|s_i; \rho).$$

[14] There is an implicit assumption in the setup that player i can directly learn only about θ_i. We make this assumption explicit in (A.11).

In words, $a_i^\star(s_i;\rho)$ is the action player i takes in an equilibrium of the game $\mathcal{G}_\rho$ when she observes signal s_i and her opponents use strategies $a_{-i}^\star(\rho)$. Fixing the basic game G, we are interested in how a change in the information structure from $\Sigma_{\rho'}$ to $\Sigma_{\rho''}$ affects the BNEs of the full game $\mathcal{G}_{\rho'} = (\Sigma_{\rho'}, G)$ and $\mathcal{G}_{\rho''} = (\Sigma_{\rho''}, G)$.

We restrict our attention to monotone BNEs, i.e., each player's equilibrium action, $a_i^\star(s_i;\rho)$ is increasing in the signal s_i.[15] The existence of monotone pure strategy BNE has long been established by the literature on supermodular Bayesian games. In particular, the existence result of Van Zandt and Vives (2007) is noteworthy in our setting; their existence result does not require players to have atomless beliefs when they participate in the basic game G, which is relevant in our setting as we do not impose any smoothness restriction on the joint distribution of signals and state.

3.2 Monotone Information Order and Equilibrium Responsiveness

We parallel the single-agent framework as closely as possible. We first extend the responsiveness order and the monotone information order into a multi-player setting using the product order. We then identify the class of payoff functions for which monotone BNE are ordered by responsiveness when information quality changes according to the monotone information order.

Equilibrium Responsiveness: Given two games of incomplete information $\mathcal{G}_{\rho'}$ and $\mathcal{G}_{\rho''}$, $a^\star(\rho'')$ is more responsive with a higher [lower] mean than $a^\star(\rho')$ if, and only if, $a_i^\star(\rho'')$ is more responsive with a higher [lower] mean than $a_i^\star(\rho')$ for all $i \in N$.

Monotone Information Order: Given two information structures $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$, $\Sigma_{\rho''}$ dominates $\Sigma_{\rho'}$ in the monotone information order, denoted $\rho'' \succeq_{MIO} \rho'$ if, and only if, $\Sigma_{\rho''_i}$ dominates $\Sigma_{\rho'_i}$ in the monotone information order for all $i \in N$.

Let $\Gamma^\uparrow$ be the class of payoff functions $u : \Theta_i \times A \to \mathbb{R}$ that satisfy (A.7)-(A.10) and have a marginal utility $u^i_{a_i}(\theta, a)$ that, for all $j \in N$,

i. is convex in a_j for all $(\theta_i, a_{-j}) \in \Theta_i \times A_{-j}$,

ii. has increasing differences in in $(\theta_i, a_{-j}; a_j)$.

[15] By assumptions (A.6), (A.10), and (A.12), player i's best response is monotone in s_i when her opponents use monotone strategies. While restricting attention to monotone BNEs may be with loss of generality, extremal equilibria are nonetheless monotone. Specifically, the least and the greatest pure strategy monotone BNEs of a supermodular Bayesian game bound all other BNEs (Milgrom and Roberts, 1990; Van Zandt and Vives, 2007).

Below, we show that payoffs in $\Gamma^{\uparrow}$ are linked to BNE strategies that become more responsive with a higher mean (hence the up arrow) as information quality increases in the monotone information order. Similar to the single agent setting, there are increasing complementarities in (θ_i, a_i) along with a marginal utility that diminishes at a diminishing rate as a_i increases. However, in the multi-player setting, player i faces uncertainty not only from θ_i but also from the random actions of her opponents. Hence, we additionally require increasing complementarities in (θ_i, a_{-i}) as well as increasing strategic complementarities in (a_i, a_{-i}) while maintaining a marginal utility that diminishes at a diminishing rate as a_{-i} increases.

Let $\Gamma^{\downarrow}$ be the class of payoff functions $u : \Theta_i \times A \to \mathbb{R}$ that satisfy (A.7)-(A.10) and have a marginal utility $u^i_{a_i}(\theta, a)$ that, for all $j \in N$,

i. is convcave in a_j for all $(\theta_i, a_{-j}) \in \Theta_i \times A_{-j}$,

ii. has decreasing differences in $(\theta_i, a_{-j}; a_j)$.

Below, we show that payoffs in $\Gamma^{\downarrow}$ are linked to BNE strategies that become more responsive with a lower mean (hence the down arrow) as information quality increases in the monotone information order.

Theorem 2 *Consider two Bayesian games, $\mathcal{G}_{\rho'} \triangleq (\Sigma_{\rho'}, G)$ and $\mathcal{G}_{\rho''} \triangleq (\Sigma_{\rho''}, G)$ in which $\Sigma_{\rho''}$ dominates $\Sigma_{\rho'}$ in the monotone information order.*

i. Suppose for each player $i \in N$, $u^i \in \Gamma^{\uparrow}$. Then, for any monotone Bayesian Nash equilibrium $a^\star(\rho')$ of $\mathcal{G}_{\rho'}$, there exists a monotone Bayesian Nash equilibrium $a^\star(\rho'')$ of $\mathcal{G}_{\rho''}$ such that $a^\star(\rho'')$ is more responsive with higher mean than $a^\star(\rho')$.

ii. Suppose for each player $i \in N$, $u^i \in \Gamma^{\downarrow}$. Then, for any monotone Bayesian Nash equilibrium $a^\star(\rho'')$ of $\mathcal{G}_{\rho''}$, there exists a monotone Bayesian Nash equilibrium $a^\star(\rho')$ of $\mathcal{G}_{\rho'}$ such that $a^\star(\rho'')$ is more responsive with lower mean than $a^\star(\rho')$.

Each player i faces n sources of uncertainty: the unknown state θ_i and the random actions of the remaining $n-1$ players which depend on the signal realizations they observe. The proof for Theorem 2 proceeds in four steps. The first step shows that, holding all else fixed, player i's best-reply strategy becomes more responsive when only the quality of information for player i increases in the monotone information order. As quality of information increases, player i is more informed about θ_i. Thus, an application of Theorem 1 from the single-agent setting gives the result.

The second step shows that, holding all else fixed, player i's best-reply strategy becomes more responsive when the quality of information for player $j \neq i$ increases in the monotone information order. As player j's information quality increases, player j's signals become more correlated to the state θ_j, which in turn is (weakly) correlated to θ_i.[16] Thus, by increasing the quality of information for player j, the signals for player i and j indirectly become more correlated. Hence, player i can better predict player j's random action and match it.

The third step shows that, holding all else fixed, player i's best-reply strategy becomes more responsive when player $j \neq i$ chooses a more responsive strategy due to the increasing differences of $u^i_{a_i}$ in $(a_j; a_i)$. It is of similar spirit to the result that increasing differences of u^i in $(a_j; a_i)$ imply that the best-reply to a monotone opponent's strategy is monotone. Finally, we conclude by applying the main result in Villas-Boas (1997) to get a comparative statics of responsiveness on fixed points.

While the requirements placed on payoff functions may be rather restrictive, we present some examples of applications in which they are satisfied.

Example 4: Beauty Contest Games

Each player's payoff is given by

$$u^i(\theta_i, a) = -\beta_i \underbrace{\left(\theta_i - a_i\right)^2}_{\text{Match the state}} - (1-\beta_i) \underbrace{\left(\sum_{j \neq i} \frac{a_j}{n-1} - a_i\right)^2}_{\text{Match average action of others}} .$$

Then, $u^i \in \Gamma^{\uparrow}$ if $\beta_i \in (0,1)$.

Example 5: Joint Projects

n players work on joint risky project which has a return of $v(\theta_i)$ to player i if it succeeds. Each player's payoff is given by

$$u^i(\theta_i, a) = \underbrace{\prod_{j=1}^{n} a_j}_{\text{Probability of success}} \times \underbrace{v(\theta_i)}_{\text{Value from success}} - \underbrace{c_i(a_i)}_{\text{Cost of effort}}$$

Then, $u^i \in \Gamma^{\uparrow}$ if

i. $v(\theta_i)$ is increasing θ_i

ii. $c_i(a_i)$ is increasing and convex in a_i, and

[16] By weakly correlated, we mean that we allow for θ_i to be independent of θ_j.

iii. $c_i'(a_i)$ is concave in a_i.

The only additional restriction than the standard assumptions made in most applications is the concavity of the marginal cost (which is satisfied if the player has quadratic cost).

Remark: In the case of independent private values, we can relax the restriction on $\Gamma^{\uparrow}$ to the class of utility functions $u^i : \Theta_i \times A \to \mathbb{R}$ which satisfy (A.7)-(A.10) and have a marginal utility $u^i_{a_i}(\theta_i, a)$ that

i. for all $j \in N$, is convex in a_j for all $(\theta_i, a_{-j}) \in \Theta_i \times A_{-j}$, and

ii. has increasing differences in $(\theta_i, a_{-i}; a_i)$.

In games with independent private values (IPV), we do not require the strategic complementarity between players to increase when the state increases. Essentially, in IPV settings, none of the information player i acquires from her signal is informative of her opponents' actions. Hence, we can drop the increasing differences assumption in $(\theta_i, a_{-j}; a_j)$ for $j \neq i$.

3.3 Application: Information Sharing in Joint Projects

Two players collaborate on a joint project which has an intrinsic common value $\theta \in \Theta \triangleq [\underline{\theta}, \bar{\theta}]$ if it succeeds. The value of a failed project is normalized to zero. Each player $i = 1, 2$ chooses an effort level $a_i \in [0, 1]$ and gets a payoff of

$$u^i(\theta, a) = a_1 a_2 v^i(\theta) - c^i(a_i)$$

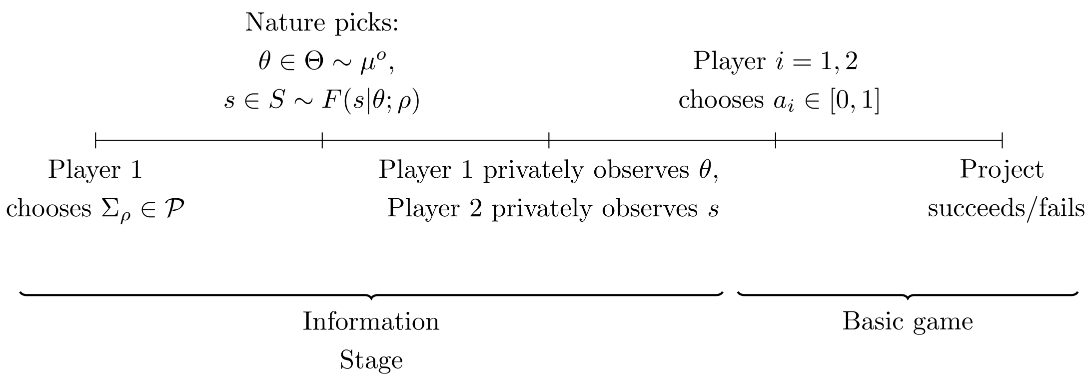
where $v^i : \Theta \to \mathbb{R}_+$ is player i's private utility from a successful project with $v^i(\cdot)$ increasing in θ, and $c^i : A_i \to \mathbb{R}_+$ is player i's cost function. We assume that $u^i \in \Gamma^{\uparrow}$ as in Example 5.

Ex-ante, the intrinsic value of the project is unknown and the players have a common prior $\mu^o \in \Delta(\Theta)$. Prior to participating in the joint project, player 1 observes the value of θ but player 2 does not.[17] Player 1 can share her information with player 2 by choosing an information structure Σ_ρ. Player 2 observes player 1's choice of information structure while player 1 cannot observe player 2's signal realization.

[17] The model in which player 1 observes θ is a special case of a more general model where player 1 only observes a signal s_1 which is correlated to θ. Then, we can treat θ as a latent variable while s_1 becomes the relevant "state of the world" with payoffs given by $\tilde{u}^i(s_1, a) = a_1 a_2 E[v^i(\theta)|s_1] - c^i \frac{a_i^2}{2}$.

Let $\mathcal{P}$ be the set of information structures from which player 1 chooses. We assume that all information structures in $\mathcal{P}$ satisfy A.11-A.13 and that for any two information structures $\Sigma_{\rho'}$, $\Sigma_{\rho''} \in \mathcal{P}$, either $\rho' \succeq_{MIO} \rho''$ or vice versa. Let $\Sigma_{\bar{\rho}} \in \mathcal{P}$ represent the full-information structure.

Each choice of an information structure Σ_ρ defines a Bayesian game $\mathcal{G}_\rho$ as follows:



Let $a^\star(\rho)$ be a monotone BNE of $\mathcal{G}_\rho$ with $a_1^\star(\cdot;\rho) : \Theta \to A_1$ and $a_2^\star(\cdot;\rho) : S \to A_2$. We assume that the players can coordinate on the maximal BNE (the BNE that results in the maximal effort for each player). Let player 1's ex-ante BNE payoff be given by

$$U_1(\rho) = \int_{\Theta \times S} u^1\big(\theta, a_1^\star(\theta;\rho), a_2^\star(s;\rho)\big) dF(s|\theta;\rho)\mu^o(d\theta).$$

How much information, if any, would player 1 want to share with player 2? On the one hand, player 1 would like to reveal the state to player 2 if θ is high as this would encourage player 2 to exert effort. In contrast, player 1 would prefer to reveal nothing to player 2 if θ is low as this would discourage player 2 from exerting effort. Ex-ante, it is not immediately clear which of these two forces dominates. However, we can use the comparative statics developed in Theorem 2 to show that it is optimal (within the restricted set of information structures) for player 1 to reveal the state to player 2.

Proposition 3
It is optimal for player 1 to reveal the state to player 2. Specifically, $U_1(\bar{\rho}) \geq U_1(\rho)$ for all $\Sigma_\rho \in \mathcal{P}$, where $\Sigma_{\bar{\rho}} \in \mathcal{P}$ is the full-information structure.

The model of joint projects with information sharing is akin to the literature on firm competition and information sharing which has been explored by a vast literature starting with Novshek and Sonnenschein (1982), Clarke (1983), Vives (1984), Gal-Or (1985, 1986), and Raith (1996). More recently, Bergemann and Morris (2013) study information sharing in beauty contests and provide a comprehensive analysis which allows for Bayes correlated equilibrium. Similar to Proposition 3, full information disclosure has been shown to be optimal for the case of firm competition with common values and strategic complements. However, the previous results depended on symmetric linear best-response functions and normally distributed states and signals so that payoffs and actions could be explicitly computed. Our result in Proposition 3 makes fewer assumptions on payoffs and information structures as we do not need explicit solutions.

4 Overt vs Covert Information Acquisition and the Value of Transparency

In this section, we study information acquisition in the context of supermodular Bayesian games. Specifically, we study how the incentives to acquire information change when information acquisition is public (overt) versus when it is private (covert) which allows us to study the strategic value of acquiring information. We call this strategic element of information acquisition the *value of transparency* and show that responsiveness is one of the key components useful in characterizing it.

4.1 Setup

We consider a two-player Bayesian game composed of two stages: an information acquisition stage followed by a basic game $G \triangleq (\{A_i, u^i\}_{i=1,2}, \mu^o)$ where $\mu^o \in \Delta(\Theta)$ satisfies (A.6) and u^i satisfies (A.7)-(A.10) for $i = 1, 2$. In the information acquisition stage, player 2 has an exogenously given information structure Σ_{ρ_2}. On the other hand, player 1 is allowed to choose an information structure $\Sigma_{\rho_1} \in \mathcal{P}$ where, similar to Section 3.3, $\mathcal{P}$ denotes the set of information structures. Again, we assume that information structures satisfy (A.11)-(A.13), and that for any two information structures $\Sigma_{\rho_1}, \Sigma_{\rho_1'} \in \mathcal{P}$, either $\rho_1 \succeq_{MIO} \rho_1'$ or vice versa. We allow information acquisition to be costly where we denote the cost of acquiring Σ_{ρ_1} by $\kappa(\rho_1) \in \mathbb{R}$.

Whether or not player 2 observes player 1's choice of information structure corresponds to the overt and the covert game respectively. After the information acquisition stage, each player $i = 1, 2$ privately observes a signal realization $s_i \in S_i$, updates beliefs, and plays the basic game by simultaneously choosing an action $a_i \in A_i$. Throughout this section, we only consider pure strategy information acquisition strategies in the first stage.[18] We also assume that players coordinate on the maximal pure-strategy monotone BNE in the second stage.

To better understand the difference between overt and covert information acquisition, suppose initially that player 1 is endowed with information structure $\Sigma_{\hat{\rho}_1} \in \mathcal{P}$ and this is common knowledge, i.e., both players know the Bayesian game is $\mathcal{G}_{\hat{\rho}} = (\Sigma_{\hat{\rho}_1}, \Sigma_{\rho_2}, G)$. Let $\big(a_1^\star(\hat{\rho}), a_2^\star(\hat{\rho})\big)$ be the resulting BNE of $\mathcal{G}_{\hat{\rho}}$. Consider the following two scenarios as a thought experiment.

In the first scenario, player 1 is allowed to choose a different information structure. Player 2 is (a) made aware that player 1 can choose a different information structure, and (b) observes player 1's choice. This scenario of the thought experiment mirrors the the overt game. Common knowledge of information structures still holds; if player 1 chooses $\Sigma_{\rho_1} \in \mathcal{P}$, the game changes from $\mathcal{G}_{\hat{\rho}}$ to $\mathcal{G}_{\rho} = (\Sigma_{\rho_1}, \Sigma_{\rho_2}, G)$ and the resulting BNE is $\big(a_1^\star(\rho), a_2^\star(\rho)\big)$.

In the second scenario, player 1 is again allowed to choose a different information structure. However, player 2 is (a) not aware that player 1 can choose a different information structure, and (b) does not observe player 1's choice. This scenario of the thought experiment mirrors the covert game. Player 2 will ignorantly believe that the game is still $\mathcal{G}_{\hat{\rho}}$ and continues to play $a_2^\star(\hat{\rho})$, even when player 1 chooses Σ_{ρ_1}. On the other hand, player 1 best-replies to $a_2^\star(\hat{\rho})$ by playing the strategy $a_1^{BR}(a_2^\star(\hat{\rho}), \rho)$. Henceforth, we refer to $\rho = (\rho_1, \rho_2)$ as the actual outcome of the information acquisition stage and $\hat{\rho} = (\hat{\rho}_1, \rho_2)$ as player 2's belief of the outcome of the information acquisition stage, i.e., player 1's actual choice of information structure is $\Sigma_{\rho_1} \in \mathcal{P}$ while player 2 believes it is $\Sigma_{\hat{\rho}_1} \in \mathcal{P}$. We say player 2 has correct beliefs when $\hat{\rho}_1 = \rho_1$ (which must be the case in any equilibrium).

Given actual first stage outcome ρ and player 2's belief $\hat{\rho}$, let player 1's ex-ante payoff in the covert game (second scenario) be $U_1(\rho; \hat{\rho}) - \kappa(\rho_1)$ where

$$U_1(\rho; \hat{\rho}) = \int_{\Theta \times S} u^1\big(\theta_1, a_1^{BR}(s_1; a_2^\star(\hat{\rho}), \rho), a_2^\star(s_2; \hat{\rho})\big) d\boldsymbol{F}(\theta, s; \rho).$$

In the overt game (first scenario), player 2 has correct beliefs. Hence, given actual first stage

[18] For overt information acquisition, this is without loss as player 2 observes the outcome of the mixed strategy before the second stage. Hence, player 1 randomizes only when she is indifferent.

outcome ρ, player 1's payoff in the overt game is $U_1(\rho;\rho) - \kappa(\rho_1)$ with

$$U_1(\rho;\rho) = \int_{\Theta\times S} u^1\big(\theta_1, a_1^{BR}(s_1; a_2^\star(\rho), \rho), a_2^\star(s_2;\rho)\big) d\boldsymbol{F}(\theta, s; \rho)$$

$$= \int_{\Theta\times S} u^1\big(\theta_1, a_1^\star(s_1;\rho), a_2^\star(s_2;\rho)\big) d\boldsymbol{F}(\theta, s; \rho),$$

where the equality follows from $a_1^{BR}(a_2^\star(\rho), \rho) = a_1^\star(\rho)$ by the definition of a BNE. We define the *value of transparency* as the difference in player 1's ex-ante payoffs between the overt and covert game.

Value of Transparency: Given actual first stage outcome ρ and player 2's belief $\hat{\rho}$, the *value of transparency* is given by:

$$VT(\rho;\hat{\rho}) = U_1(\rho;\rho) - U_1(\rho;\hat{\rho}).$$

In words, $VT(\rho;\hat{\rho})$ represents the gain/loss to player 1 from disclosing to player 2 her actual first stage choice, Σ_{ρ_1}, instead of letting player 2 incorrectly believe that the first stage choice is $\Sigma_{\hat{\rho}_1}$. Notice that the value of transparency does not capture any direct substantive advantages of information; player 1's chosen information structure in both cases is Σ_{ρ_1}. The difference between covert and overt stems from player 2's beliefs about player 1's quality of information.

4.2 Demand for Information in Overt vs Covert Games

Before we discuss how to characterize the value of transparency, we present why it is an interesting economic concept. In particular, we show that the value of transparency is helpful in answering the following questions: when is a higher quality of costless but overt information acquisition always beneficial to player 1? Does player 1 acquire more information when information acquisition is overt or when it is covert?

In covert games, the more information player 1 acquires, the more knowledgeable she is about the unknown state and can make better decisions in the game. Therefore, if information is costless, the value of a higher quality information in covert games is positive (Neyman; 1989, Amir and Lazzati; 2016).

While acquiring a higher quality of information has the same positive effect in overt games, there are additional effects to account for; player 2 can observe how much information player

1 acquires and respond to it during the basic game. Player 2 may find it optimal to choose an unfavorable action (punish player 1) in the equilibrium of the second stage whenever player 1 acquires more information in the first stage. If player 2's unfavorable action is strong enough on average, player 1's value of a higher quality information in overt games may be negative despite becoming more informed.

Proposition 4 below states that this is not the case if the value of transparency is non-negative whenever player 1 acquires an information structure of higher quality than player 2's belief. In other words, player 1 benefits by disclosing to player 2 that the actual choice of information in the first stage whenever the actual choice is of higher quality than player 2's initial belief.

Proposition 4 *Let κ be a constant function. For any two information structures $\Sigma_{\rho_1}, \Sigma_{\hat{\rho}_1} \in \mathcal{P}$, suppose $\rho_1 \succeq_{MIO} \hat{\rho}_1$ implies $VT(\rho; \hat{\rho}) \geq 0$. Then, $U_1(\rho; \rho) \geq U_1(\hat{\rho}; \hat{\rho})$.*

Intuitively, player 1's payoff difference between overtly acquiring Σ_{ρ_1} versus $\Sigma_{\hat{\rho}_1}$ is given by

$$\underbrace{U_1(\rho;\rho) - U_1(\hat{\rho};\hat{\rho})}_{\text{value of overt information}} = \underbrace{U_1(\rho;\rho) - U_1(\rho;\hat{\rho})}_{\text{value of transparency}} + \underbrace{U_1(\rho;\hat{\rho}) - U_1(\hat{\rho};\hat{\rho})}_{\text{value of covert information}}.$$

The above expression decomposes the value of overtly acquiring a higher quality of information into the value of covertly acquiring a higher quality of information and the value from disclosing to player 2 that a higher quality of information has been acquired. The latter effect is the value of transparency. The value of covert information captures the change in player 1's payoff when the quality of information increases from $\Sigma_{\hat{\rho}_1}$ to Σ_{ρ_1} while player 2's beliefs are held fixed at $\Sigma_{\hat{\rho}_1}$. Amir and Lazzati (2016) show that the value of covert information in supermodular games is non-negative whenever $\rho_1 \succeq_{MIO} \hat{\rho}_1$. Hence, if the value of transparency is also non-negative whenever information quality increases in the monotone information order, then the value of overt information is non-negative.

To answer the second question about the demand of information, let $\Sigma_{\rho_1^c}$ and $\Sigma_{\rho_1^o}$ denote the information structures acquired in equilibrium under covert and overt games. Specifically, $\Sigma_{\rho_1^c}$ is a solution to

$$\max_{\Sigma_{\rho_1} \in \mathcal{P}} U_1(\rho; \rho^c) - \kappa(\rho_1).$$

In words, given player 2 believes player 1 chooses $\Sigma_{\rho_1^c}$ in equilibrium, it is indeed optimal for player 1 to choose $\Sigma_{\rho_1^c}$. In contrast, $\Sigma_{\rho_1^o}$ solves

$$\max_{\Sigma_{\rho_1}\in\mathcal{P}} U_1(\rho;\rho) - \kappa(\rho_1).$$

In words, $\Sigma_{\rho_1^o}$ is optimal for player 1 after taking into account that player 2 will observe the chosen information structure in the first stage and will respond to it in the second stage. Proposition 5 below show that whenever the value of transparency is non-negative, player 1 acquires more information in overt games than in covert games, regardless of the cost function.

Proposition 5 *For any two information structures $\Sigma_{\rho_1}, \Sigma_{\hat{\rho}_1} \in \mathcal{P}$, suppose $VT(\rho;\hat{\rho}) \geq 0$ if, and only if, $\rho_1 \succeq_{MIO} \hat{\rho}_1$. Then, $\rho_1^o \succeq_{MIO} \rho_1^c$.*

The implicit assumption of a unique equilibrium outcomes in the result above is only made to simplify exposition. Notice that the antecedent of Proposition 5 implies VT is a single crossing function, i.e., $VT(\hat{\rho};\hat{\rho}) = 0$ and $VT(\rho;\hat{\rho}) \geq 0$ for any $\rho_1 \succeq_{MIO} \hat{\rho}_1$. We can therefore apply familiar monotone comparative statics tools to show that the solution set for overt equilibrium maximization problem dominates the solution set for covert equilibrium in the monotone-information order.

4.3 Characterizing the Value of Transparency

We now discuss how to characterize the value of transparency. We show that the value of transparency depends on ($i.$) the responsiveness of player 2's equilibrium action to changes in the quality of player 1's information structure, and ($ii.$) the externality player 2's response imposes on player 1.

Theorem 3 *Assume either*

i. states are independently distributed across players, or

ii. $u^1(\theta_1, a)$ has increasing differences in $(\theta_1; a_2)$.

Furthermore, suppose $u^1(\theta_1, a)$ is an increasing convex [decreasing concave] function of a_2, and $a_2^\star(\cdot)$ becomes more responsive with a higher [lower] mean as information quality increases in the monotone information order. Then, for any two information structures $\Sigma_{\rho_1}, \Sigma_{\hat{\rho}_1} \in \mathcal{P}$, $VT(\rho;\hat{\rho}) \geq 0$ if, and only if, $\rho_1 \succeq_{MIO} \hat{\rho}_1$.

Notice that both the beauty contest game and the joint project game presented in Example 4 and Example 5 of Section 3 satisfy the conditions of Theorem 3. Applying Proposition 5, we can then conclude that the demand for information in a beauty contest or joint project game is higher when information acquisition is overt.

To gain some intuition for Theorem 3, recall that $VT(\rho;\hat{\rho}) = U_1(\rho;\rho) - U_1(\rho;\hat{\rho})$ is given by

$$\int_{\Theta\times S}\Big[u^1\big(\theta_1, a_1^\star(s_1;\rho), a_2^\star(s_2;\rho)\big) - u^1\big(\theta_1, a_1^{BR}(s_1; a_2^\star(\hat{\rho}),\rho), a_2^\star(s_2;\hat{\rho})\big)\Big]d\boldsymbol{F}(\theta, s;\rho).$$

By taking a first-order Taylor expansion, we can approximate the value of transparency as

$$\approx \int_{S_1}\Big(a_1^\star(s_1;\rho) - a_1^{BR}(s_1; a_2^\star(\hat{\rho}),\rho)\Big)\underbrace{\int_{\Theta\times S_2} u^1_{a_1}\big(\theta_1, a_1^\star(s_1;\rho), a_2^\star(s_2;\rho)\big)d\boldsymbol{F}(\theta, s_2|s_1;\rho)}_{\text{=0 by first-order condition for optimality}}\,dF_{S_1}(s_1)$$

$$+\int_{S_2}\Big(a_2^\star(s_2;\rho) - a_2^\star(s_2;\hat{\rho})\Big)\int_{\Theta\times S_1} u^1_{a_2}\big(\theta_1, a_1^\star(s_1;\rho), a_2^\star(s_2;\rho)\big)d\boldsymbol{F}(\theta, s_1|s_2;\rho)dF_{S_2}(s_2).$$

The second term is the interaction of player 2's responsiveness to changes in player 1's information quality, captured by the difference in $a_2^\star(\rho)$ and $a_2^\star(\hat{\rho})$, and the effect (externality) this responsiveness has on player 1's expected payoff, captured by $u^1_{a_2}$. The conditions in Theorem 3 allow us to sign the second expression of the Taylor expansion. For example, if states are independently distributed across players,

$$\int_{S_2}\Big(a_2^\star(s_2;\rho) - a_2^\star(s_2;\hat{\rho})\Big)\int_{\Theta\times S_1} u^1_{a_2}\big(\theta_1, a_1^\star(s_1;\rho), a_2^\star(s_2;\rho)\big)d\boldsymbol{F}(\theta, s_1|s_2;\rho)dF_{S_2}(s_2)$$

$$= \int_{S_2}\Big(a_2^\star(s_2;\rho) - a_2^\star(s_2;\hat{\rho})\Big)dF_{S_2}(s_2)\int_{\Theta_1\times S_1} u^1_{a_2}\big(\theta_1, a_1^\star(s_1;\rho), a_2^\star(s_2;\rho)\big)dF(\theta_1, s_1;\rho_1).$$

If $a_2^\star$ becomes more responsive with a higher mean when information quality increases, then

$$\int_{S_2}\Big(a_2^\star(s_2;\rho) - a_2^\star(s_2;\hat{\rho})\Big)dF_{S_2}(s_2) \geq 0.$$

The sign for the value of transparency is then pinned down by whether $u^1_{a_2} \geq 0$ (positive externalities) or $u^1_{a_2} \leq 0$ (negative externalities).

4.4 Relation to Strategic Effects of Investment in Firm Competition

The application of responsiveness to characterize the value of transparency is related to the taxonomy of strategic behavior in firm competition studied by Fudenberg and Tirole (1984), and Bulow, Geanakoplos and Klemperer (1985).[19] Here we follow the textbook treatment of Tirole (1988) and only consider the case of entry accommodation in a duopoly under complete information.

There are two periods and two firms, an incumbent (firm 1) and an entrant (firm 2). In the first period, the incumbent chooses a level of investment $K_1 \in \mathbb{R}$, which the entrant observes. The term investment is used in a very broad sense and can represent, for example, investment in R&D that lowers the incumbent's marginal costs or advertising that captures a share of the market.

In the second period, both firms compete either in quantities (strategic substitutes) or prices (strategic complements). Let $\big(a_1^\star(K_1), a_2^\star(K_1)\big)$ be the resulting Nash equilibrium of the second period after the incumbent chose K_1 in the first period. The incumbent's payoff from choosing an investment level K_1 is given by $U_1\big(K_1, a_1^\star(K_1), a_2^\star(K_1)\big)$.

Fudenberg and Tirole (1984) show that the total marginal effect on the incumbent's payoff from overtly increasing investment can be decomposed into

$$\frac{dU_1}{dK_1} = \underbrace{\underbrace{\frac{\partial U_1}{\partial K_1}}_{\text{direct effect}} + \underbrace{\frac{\partial U_1}{\partial a_1}\frac{da_1^\star}{dK_1}}_{\substack{=0 \\ \text{by Envelope theorem}}}}_{\text{value of covert investment}} + \underbrace{\frac{\partial U_1}{\partial a_2}\frac{da_2^\star}{dK_1}}_{\text{strategic effect}}.$$

Increasing the level of investment has a direct effect on the incumbent's payoff, for example, by reducing the marginal cost. It also affects the incumbent's optimal action choice in the second period, captured by $\frac{da_1^\star}{dK_1}$. If the entrant was unable to observe the incumbent's investment choice, these would be the only marginal effects to account for when the incumbent increases investment.

However, since the entrant observes the incumbent's first period choice of K_1, the investment also has strategic effects; the entrant's production/pricing decision is indirectly affected by K_1. This strategic effect depends on how the entrant's equilibrium strategy responds to an increase in the level of investment, represented by $\frac{da_2^\star}{dK_1}$, and on how the entrant's actions affect the incumbent's payoff. The latter effect is the externality, $\frac{\partial U^1}{\partial a_2}$, the entrant imposes on the

[19]For a thorough treatment of different examples and applications we recommend Shapiro (1986). For a more recent treatment using the tools of supermodular games see Vives (2000).

incumbent.

In our model, the game is one of incomplete information, player 1 is the incumbent, player 2 is the entrant, and the investment level K_1 corresponds to the quality of the player 1's information structure ρ_1. The total effect of increasing overt investment in information from Σ_{ρ_1} to $\Sigma_{\hat{\rho}_1}$ can be similarly decomposed into

$$U_1(\rho;\rho) - U_1(\hat{\rho};\hat{\rho}) = \underbrace{U_1(\rho;\hat{\rho}) - U_1(\hat{\rho};\hat{\rho})}_{\text{value of covert investment}} + \underbrace{U_1(\rho;\rho) - U_1(\rho;\hat{\rho})}_{\text{strategic effect}}.$$

The value of covert investment (value of covert information) captures how the player 1's payoff increase by her ability to make better informed decisions while holding the player 2's strategy fixed. The strategic effect in our model corresponds to the value of transparency. It captures how player 1's payoff changes when the player 2's strategy is indirectly affected by the change in information quality. From the first-order Taylor expansion, we have shown that the strategic effect of information (value of transparency) depends on the responsiveness of player 2's equilibrium strategy, $a_2^\star(\rho)$, and the average externalities imposed on player 1 by player 2's responsiveness. Hence, our characterization of the value of transparency can be thought of as a stochastic extension to the characterization of strategic effects of investment by Fudenberg and Tirole (1984). Generally, characterizing strategic effects of investment in a stochastic environment is a more involved exercise as it requires characterizing the change to the entire distribution of equilibrium outcomes as investment increases. However, we were able to overcome these difficulties by applying the comparative statics we developed in Section 3 for supermodular games.

5 Conclusion

In this paper, we add to the literature on comparative statics in several directions. First, we conceptualize the ex-ante perspective in which the optimal actions of a decision maker are endogenous random variables. A natural question then is how to compare optimal actions as the quality of the information changes. We introduce the notion of responsiveness and characterize how it captures changes in the mean and dispersion of actions.

We show that when payoffs exhibit increasing supermodularity, optimal actions become more responsive when the quality of information increases. The conditions can be interpreted as measuring the relative strength of the complementarity between actions and states. Furthermore, we extend our results to games of incomplete information with strategic complementarities to show that monotone Bayesian Nash equilibrium strategies become more responsive when the quality of information increases for at least one player.

In the final section we introduce the value of transparency which captures the strategic effects of information in overt (public) vs covert (private) information acquisition games. We show that the concept of responsiveness is a useful tool to characterize the value of transparency which, in turn, has implications on the value and demand for information when information acquisition is overt.

We expect that the methods of 'ex-ante comparative statics' will be specially fruitful when studying Bayesian persuasion with restricted information structures, signal jamming games, rational inattention, firm competition, and search.

References

Amir, R. and Lazzati, N. . Endogenous information acquisition in bayesian games with strategic complementarities. *Journal of Economic Theory*, 163:684–698, 2016.

Athey, S. . Monotone comparative statics under uncertainty. *The Quarterly Journal of Economics*, 117(1):187–223, 2002.

Athey, S. and Levin, J. . The value of information in monotone decision problems. *Research in Economics*, 2017. ISSN 1090-9443.

Blackwell, D. . Comparison of experiments. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 93–102, Berkeley, Calif., 1951. University of California Press.

Blackwell, D. . Equivalent comparisons of experiments. *Ann. Math. Statist.*, 24(2):265–272, 06 1953.

Bulow, J. I. , Geanakoplos, J. D. , and Klemperer, P. D. . Multimarket oligopoly: Strategic substitutes and complements. *Journal of Political economy*, 93(3):488–511, 1985.

Fudenberg, D. and Tirole, J. . The fat-cat effect, the puppy-dog ploy, and the lean and hungry look. *The American Economic Review*, 74(2):361–366, 1984.

Jensen, M. K. . Distributional comparative statics. *The Review of Economic Studies*, 85(1): 581–610, 2018.

Jewitt, I. . Information order in decision and agency problems. 01 2006.

Kamenica, E. and Gentzkow, M. . Bayesian persuasion. *American Economic Review*, 101(6): 2590–2615, 2011.

Lehmann, E. L. . Comparing location experiments. *Ann. Statist.*, 16(2):521–533, 06 1988.

Lu, J. . Random choice and private information. *Econometrica*, 84(6):1983–2027, 2016. ISSN 1468-0262.

Mas-Colell, A. , Whinston, M. D. , Green, J. R. , et al. *Microeconomic theory*, volume 1. Oxford university press New York, 1995.

Milgrom, P. and Roberts, J. . Comparing equilibria. *The American Economic Review*, pages 441–459, 1994.

Persico, N. . Information acquisition in auctions. *Econometrica*, 68(1):135–148, 2000. ISSN 1468-0262.

Quah, J. K.-H. and Strulovici, B. . Comparative statics, informativeness, and the interval dominance order. *Econometrica*, 77(6):1949–1992, 2009. ISSN 1468-0262.

Shaked, M. and Shanthikumar, J. G. . *Stochastic orders*. Springer Science & Business Media, 2007.

Van Zandt, T. and Vives, X. . Monotone equilibria in bayesian games of strategic complementarities. *Journal of Economic Theory*, 134(1):339–360, 2007.

Villas-Boas, J. M. . Comparative statics of fixed points. *Journal of Economic Theory*, 73(1): 183–198, 1997.

Vives, X. . Nash equilibrium with strategic complementarities. *Journal of Mathematical Economics*, 19(3):305–321, 1990.

REFERENCES

AMIR, R. and N. LAZZATI (2016): "Endogenous information acquisition in Bayesian games with strategic complementarities," *Journal of Economic Theory*, 163, 684-698.

ATHEY, S (2002): "Monotone Comparative Statics under Uncertainty," *Quarterly Journal of Economics*, CXVII(1), 187-223.

ATHEY, S. and J. LEVIN (2017): "The Value of Information in Monotone Decision Problems," *Research in Economics* 01-003.

BLACKWELL, D. (1951): "Comparison of experiments," *Second Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, 93-102.

—— (1953): "Equivalent Comparisons of Experiments," *The Annals of Mathematical Statistics*, 24(2), 265-272.

BERGEMANN, D., and S. MORRIS (2016): "Bayes Correlated Equilibrium and the Comparison of Information Structures in Games," *Theoretical Economics*, 11, 487-522.

—— (2013):"Robust Predictions in Games with Incomplete Information," *Econometrica*, 81(4), 1251-1308.

BULOW, J. I., J. D. GEANAKOPLOS and P. D. KLEMPERER (1985): "Multimarket Oligopoly: Strategic Complements and Substitutes," *Journal of Political Economy*, 93(3), 488-511.

CLARKE, R. N. (1983): "Collusion and the Incentives for Information Sharing," *The Bell Journal of Economics*, 383-394.

COLOMBO, L., G. FEMMINIS, and A. PAVAN (2014): "Information Acquisition and Welfare," *The Review of Economic Studies*, 81, 1438-1483.

DENTI. T. (2016): "Unrestricted Information Acquisition," *Working Paper*.

FUDENBERG, D. and J. TIROLE (1984): "The Fat-Cat Effect, the Puppy-Dog Ploy, and the Lean and Hungry Look," *American Economic Review (Papers and Proceedings)*, 74(2),361-366.

GAL-OR, E. (1985): "Information Sharing in Oligopoly," *Econometrica*, 329-343.

—— (1986): "Information Transmission-Cournot and Bertrand Equilibria," *The Review of Economic Studies*, 53(1), 85-92.

GANUZA, J.J., and J. PENALVA (2010): "Signal Ordering Based on Dispersion and the Supply of Private Information in Auctions," *Econometrica*, 78 (3), 1007-1030.

HELLWIG, C. and L. VELDKAMP (2009): "Knowing What Others Know: Coordination Motives in Information Acquisition," *Review of Economic Studies*, 76, 223-251.

JENSEN, M.K. (2017): "Distributional Comparative Statics," *The Review of Economic Studies*, 10.1093.

JEWITT, I. (2007): "Information Order in Decision and Agency Problems," *mimeo*, Nuffield College.

LEHMANN, E. L. (1988): "Comparing Location Experiments," *The Annals of Statistics*, 16 (2), 521-533.

LEVIN, J. (2001): "Information and the Market for Lemons," *Rand Journal of Economics*, 32(4), 657-666.

LU, J, (2016): "Random Choice and Private Information," *Econometrica*, 84 (6), 19832027.

MEYER, M.A. (1990): "Interdependence in Multivariate Distributions: Stochastic Dominance Theorems and an Application to the Measurement of Ex Post Inequality under Uncertainty," *Nuffield College D. P.* No. 49.

MEYER, M. A., and D. MOOKHERJEE (1987): "Incentives, Compensation and Social Welfare," *The Review of Economic Studies*, 45, 209-26.

MEYER, M.A., and B. STRULOVICI (2012): "Increasing Interdependence of Multivariate Distributions," *Journal of Economic Theory*, 147, 1460-89.

MILGROM, P. (1981): "Good News and Bad News: Representation Theorems and Applications," *The Bell Journal of Economics*, 12(2), 380-391.

MILGROM, P. and J. ROBERTS (1990): "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities," *Econometrica*, 58(6), 1255-1277.
—— (1994): "Comparing Equilibria," *American Economic Review*, 84(3), 441-459.

MILGROM, P. and C. SHANNON (1994): "Monotone Comparative Statics," *Econometrica,* 62(1), 157-180.

MULLER, A., D. STOYAN (2002): "Comparison Methods for Stochastic Models and Risks," *John Wiley & Sons, Ltd.*

MYATT, D. P. and C. WALLACE (2012): "Endogenous Information Acquisition in Coordination Games," *The Review of Economic Studies*, 79, 340-374.

NEYMAN, A. (1989): "The Positive Value of Information," *Games and Economic Behavior*, 3, 350-355 (1991)

NOVSHEK, W., and SONNENSCHEIN, H. (1982): "Fulfilled Expectations Cournot Duopoly with Information Acquisition and Release," *The Bell Journal of Economics*, 214-218.

PAVAN, A. (2016): "Attention, Coordination and Bounded Recall," *Working Paper*.

PERSICO, N. (1996): "Information Acquisition in Affiliated Decision Problems," *Discussion Paper* No. 1149, Department of Economics, Northwestern University.
—— (2000): "Information Acquisition in Auctions," *Econometrica*, 68(1), 135-148.

QUAH, J. and B. STRULOVICI (2009): "Comparative Statics, Informativeness, and the Interval Dominance Order," *Econometrica*, 77(6), 1949-1992.

RAITH, M. (1996): "A General Model of Information Sharing in Oligopoly," *Journal of economic theory*, 71(1), 260-288.

SCHILLING, R. (2005): "Measures, Integrals and Martingales," *Cambridge University Press*, Cambridge CB2 8RU, UK.

SHAKED, M. and G. SHANTHIKUMAR (2007): "Stochastic Orders," *Springer-Verlag* New York, XVI, 474pp.

SHAPIRO, C. (1986): "Theories of Oligopoly Behavior," In *Handbook of Industrial Organization*, ed. R. Schmalensee and R. Willig (Amsterdam: North Holland).

TIROLE, J. (1988): "The Theory of Industrial Organization," *MIT Press*, 479pp.
—— (2015): "Cognitive Games and Cognitive Traps," *Working Paper*.

TOPKIS, D. M. (1978): "Minimizing a Submodular Function on a lattice," *Operations Research*, 26, 305-321.

VAN ZANDT, T., and X. VIVES (2007): "Monotone Equilibria in Bayesian Games of Strategic Complementarities," *Journal of Economic Theory*, 134, 339-360.

VILLAS-BOAS, J.M. (1997): "Comparative Statics of Fixed Points," *Journal of Economic Theory*, 73, 183-198.

VIVES, X. (1984): "Duopoly Information Equilibrium: Cournot and Bertrand," *Journal of Economic Theory*, 34(1), 71-94.
—— (1990): "Nash Equilibrium with Strategic Complementarities," *Journal of Mathematical Economics*, 19, 305-21.
—— (2000): "Oligopoly Pricing: Old Ideas and New Tools," *MIT Press*, 441pp.
—— (2008): "Information Sharing Among Firms," The New Palgrave Dictionary of Economics. Eds. Steven N. Durlauf and Lawrence E. Blume.

YANG, M., 2015: "Coordination with Flexible Information Acquisition," *Journal of Economic Theory*, 158(B), 721-738.

6 Appendix

6.1 Proofs from Section 2

The following alternative characterizations of the monotone information order will prove useful for the proof of Theorem 1.

Lemma 2 *Given two information structures $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$, $\rho'' \succeq_{MIO} \rho'$ if, and only if,*

i. For all $(\theta, s) \in \Theta \times S$,

$$F(\theta, s; \rho'') \geq F(\theta, s; \rho').$$

ii. For all integrable supermodular functions $\psi : \Theta \times S \to \mathbb{R}$,

$$\int_{\Theta \times S} \psi(\theta, s) dF(\theta, s; \rho'') \geq \int_{\Theta \times S} \psi(\theta, s) dF(\theta, s; \rho')$$

Proof. Fix any $(\theta, s) \in \Theta \times S$. Let $F_S(s) = q$.

$$\begin{aligned}
F(\theta, s; \rho'') - F(\theta, s; \rho') &= \int_{-\infty}^{s} \int_{\underline{\theta}}^{\theta} \Big(\mu(d\omega|x; \rho'') - \mu(d\omega|x; \rho')\Big) dF_S(x) \\
&= \left(\int_{\underline{\theta}}^{\theta} \mu(d\omega|x \leq s; \rho'') - \int_{\underline{\theta}}^{\theta} \mu(d\omega|x \leq s; \rho') \right) F_S(s) \\
&= \left(\int_{\underline{\theta}}^{\theta} \mu(d\omega|x \leq F_S^{-1}(q); \rho'') - \int_{\underline{\theta}}^{\theta} \mu(d\omega|x \leq F_S^{-1}(q); \rho') \right) q \\
&= \left(\int_{\underline{\theta}}^{\theta} \mu(d\omega|F_S(x) \leq q; \rho'') - \int_{\underline{\theta}}^{\theta} \mu(d\omega|F_S(x) \leq q; \rho') \right) q.
\end{aligned}$$

We then have

$$\begin{aligned}
\rho'' \succeq_{MIO} \rho' &\Leftrightarrow \mu(\cdot|F_S(x) \leq q; \rho') \succeq_{FOSD} \mu(\cdot|F_S(x) \leq q; \rho'') \\
&\Leftrightarrow \left(\int_{\underline{\theta}}^{\theta} \mu(d\omega|F_S(x) \leq q; \rho'') - \int_{\underline{\theta}}^{\theta} \mu(d\omega|F_S(x) \leq q; \rho') \right) q \geq 0
\end{aligned}$$

giving us the desired result. By Bayes consistency, we also have

$$\int_s^\infty \int_{\underline{\theta}}^\theta \mu(d\omega|x;\rho'')dF_S(s) \leq \int_s^\infty \int_{\underline{\theta}}^\theta \mu(d\omega|x;\rho')dF_S(s).$$

Lemma 2.*ii* follows from Müller and Stoyan (2002), Theorem 3.9.5. ■

Proof of Theorem 1

Proof.

($\Longrightarrow$) The payoff $u(\theta,a)$ is supermodular in (θ,a) and the information structure Σ_ρ has the property that $s>s'$ implies $\mu(\cdot|s;\rho)\succeq_{FOSD}\mu(\cdot|s';\rho)$. From monotone comparative statics, the optimal action $a(\rho):S\to A$ is a monotone function of s. Hence, from an ex-ante perspective, the optimal action coincdes with the quantile function we used to define responsiveness in Lemma 1, i.e., $a(\rho)=\hat{a}(\rho)$ almost surely.

Without loss of generality, we assume that the marginal on signals is uniformly distributed on the unit interval.[20] For any two information structures $\rho''\succeq_{MIO}\rho''$ and any signal realization $s\in[0,1]$, the first order conditions imply that

$$\int_\Theta u_a(\theta,a(s;\rho''))\mu(d\theta|s;\rho'') - \int_\Theta u_a(\theta,a(s;\rho'))\mu(d\theta|s;\rho') = 0$$

which we rewrite as

$$\int_\Theta \Big(u_a(\theta,a(s;\rho'')) - u_a(\theta,a(s;\rho'))\Big)\mu(d\theta|s;\rho'') + \int_\Theta u_a(\theta,a(s;\rho'))\Big(\mu(d\theta|s;\rho'')-\mu(d\theta|s;\rho')\Big) = 0$$

If $u\in\mathcal{U}^\uparrow$, then $u_a(\theta,a)$ is convex in a for all θ. Thus,

$$u_a(\theta,a(s;\rho'')) - u_a(\theta,a(s;\rho')) \geq u_{aa}(\theta,a(s;\rho'))\big(a(s;\rho'')-a(s;\rho')\big)$$

and

$$\Big(a(s;\rho'')-a(s;\rho')\Big)\int_\Theta u_{aa}(\theta,a(s;\rho'))\mu(d\theta|s;\rho'') + \int_\Theta u_a(\theta,a(s;\rho'))\Big(\mu(d\theta|s;\rho'')-\mu(d\theta|s;\rho')\Big) \leq 0.$$

[20] As mentioned in the text, we can apply the integral probability transformation to signals.

For each $t \in [0,1]$,

$$\int_t^1 \big(a(s;\rho') - a(s;\rho'')\big)ds$$

$$\leq \int_t^1 \Bigg(\underbrace{-\int_\Theta u_{aa}(\theta, a(s;\rho'))\mu(d\theta|s;\rho'')}_{\triangleq B(s)} \Bigg)^{-1} \int_\Theta u_a(\theta, a(s;\rho'))\Big(\mu(d\theta|s;\rho') - \mu(d\theta|s;\rho'')\Big)ds$$

$$= \int_{[0,1]\times\Theta} u_a(\theta, a(s;\rho'))B(s)^{-1}\mathbb{1}_{[s\geq t]}\Big(\mu(d\theta|s;\rho') - \mu(d\theta|s;\rho'')\Big)ds,$$

where $\mathbb{1}_{[s\geq t]}$ is the indicator function that equals 1 if $s \geq t$ and 0 otherwise.
Define $\psi(\theta, s; t) \triangleq u_a(\theta, a(s;\rho'))B(s)^{-1}\mathbb{1}_{[s\geq t]}$. For any $\theta > \theta'$, $\psi(\theta, s; t) - \psi(\theta', s; t) = 0$ for $s < t$ and

$$\psi(\theta, s; t) - \psi(\theta', s; t) = B(s)^{-1}\Big(u_a(\theta, a(s;\rho')) - u_a(\theta', a(s;\rho'))\Big) \geq 0$$

for $s \geq t$. The inequality follows from the supermodularity of u in (θ, a) and the strict concavity of u in a. Since $u \in \mathcal{U}^\uparrow$, u_a is also supermodular in (θ, a), i.e., $u_a(\theta, a) - u_a(\theta', a)$ is increasing in a. Since $a(s;\rho')$ is increasing in s, $u_a(\theta, a(s;\rho')) - u_a(\theta', a(s;\rho'))$ is also increasing in s. Additionally, $u \in \mathcal{U}^\uparrow$ implies that $-u_a$ is submodular in (θ, a) and concave in a. Hence, $-u_{aa}(\theta, a)$ is decreasing in both θ and a. Since higher signal realizations lead to higher actions and to first-order stochastic shifts in beliefs,

$$-\int_\Theta u_{aa}(\theta, a(s;\rho'))\mu(d\theta|s;\rho'')$$

is a decreasing function of s. Thus $B(s)^{-1}$ is increasing in s. We can therefore conclude that $\psi(\theta, s; t) - \psi(\theta', s; t)$ is increasing in s. In other words, $\psi(\theta, s; t)$ is supermodular in (θ, s). Thus, for each $t \in [0,1]$,

$$\int_t^1 \big(a(s;\rho') - a(s;\rho'')\big)ds$$

$$\leq \int_{[0,1]\times\Theta} \psi(\theta, s; t)\Big(\mu(d\theta|s;\rho') - \mu(d\theta|s;\rho'')\Big)ds \leq 0$$

where the last inequality follows from the characterization of monotone information order in Lemma 2.

($\Longleftarrow$) From Lemma 2, if $\rho'' \not\geq_{MIO} \rho'$, there exists a $(\theta^*, s^*) \in \Theta \times [0,1]$ such that

$$F(\theta^*, s^*; \rho'') < F(\theta^*, s^*; \rho').$$

Define a payoff function

$$u(\theta, a) = -\frac{1}{2}\Big(\bar{a} - \mathbb{1}_{[\theta \leq \theta^*]}(\bar{a} - \underline{a}) - a\Big)^2.$$

The payoff $u(\theta, a)$ satisfies (A.1)-(A.4): It is continuous, twice differentiable, and strictly concave in a for each $\theta \in \Theta$. It is supermodular in (θ, a). For each $\theta \in \Theta$, the optimal action is easily computed from the first order conditions so that the optimal action under complete information is $\underline{a}$ if $\theta \leq \theta^*$ and $\bar{a}$ otherwise.
Furthermore, the marginal utility $u_a(\theta, a) = \bar{a} - \mathbb{1}_{[\theta \leq \theta^*]}(\bar{a} - \underline{a}) - a$ is

i. linear in a for all $\theta \in \Theta$

ii. modular in (θ, a).

Therefore, $u \in \mathcal{U}^{\uparrow} \bigcap \mathcal{U}^{\downarrow}$.
For any given Σ_ρ, notice that

$$\begin{aligned} a(s; \rho) =& \bar{a} - (\bar{a} - \underline{a}) E\left[\mathbb{1}_{[\theta \leq \theta^*]} | s; \rho\right] \\ =& \bar{a} - (\bar{a} - \underline{a}) \int_{\underline{\theta}}^{\theta^*} \mu(d\omega | s; \rho). \end{aligned}$$

Then given $\Sigma_{\rho'}$and $\Sigma_{\rho''}$,

$$\begin{aligned} &\int_0^{s^*} \big(a(s; \rho'') - a(s; \rho')\big) dF_S(s) \\ =& (\bar{a} - \underline{a}) \Big(F(\theta^*, s^*; \rho') - F(\theta^*, s^*; \rho'')\Big) > 0. \end{aligned}$$

Therefore, $a(\rho'')$ is not more responsive with a lower mean than $a(\rho')$.
Notice that for any Σ_ρ,

$$E[a(\rho)] = \bar{a} - (\bar{a} - \underline{a}) \int_0^1 \int_{\underline{\theta}}^{\theta^*} \mu(d\omega | s; \rho) dF_S(s) = \bar{a} - (\bar{a} - \underline{a}) \int_{\underline{\theta}}^{\theta^*} \mu^o(d\omega).$$

Thus,

$$\int_{t^*}^{1} \big(a(s;\rho'') - a(s;\rho')\big) dF_S(s)$$

$$= \underbrace{\int_0^1 \big(a(s;\rho'') - a(s;\rho')\big) dF_S(s)}_{=0} - \left(\underbrace{\int_0^{t^*} \big(a(s;\rho'') - a(s;\rho')\big) dF_S(s)}_{>0} \right) < 0$$

and thus, $a(\rho'')$ is not more responsive with a higher mean than $a(\rho')$. ■

Proof of Proposition 2

Proof. $-qP''(q)/P'(q) \leq 1$ implies that $CS(q) = \int_0^q P(t)dt - qP(q)$ is an increasing convex function. If $\pi \in \mathcal{U}^{\uparrow}$, an increase in ρ (higher quality of information by MIO) leads to an optimal action $q^M(\rho)$ that is more responsive with a higher mean. By definition of responsiveness with a higher mean, $E[CS(q^M(\rho))]$ is increasing in ρ. ■

6.2 Proofs from Section 3

Proof of Theorem 2

Proof. To simplify exposition, let $n = 2$. Once again, we assume without loss of generality that for each player $i = 1, 2$, the marginal on signals, F_{S_i}, is the uniform distribution on the unit interval. Fix a Bayesian game $\mathcal{G}_\rho$. For each player i, let $\alpha_i : S_i \to A_i$ be an arbitrary measurable and monotone strategy. Let $\mathcal{A}_i$ be the set of all such monotone and measurable strategies and let $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2$. Given opponent strategies $\alpha_{-i} \in \mathcal{A}_{-i}$, let $a_i^{BR}(\alpha_{-i}, \rho) : S_i \to A_i$ be player i's best response strategy. Specifically,

$$a_i^{BR}(s_i; \alpha_{-i}, \rho) = \arg\max_{a_i \in A_i} \int_{\Theta \times \mathcal{S}_{-i}} u^i\Big(\theta_i, \alpha_{-i}(s_{-i}), a_i\Big) d\boldsymbol{F}(\theta, s_{-i}|s_i; \rho).$$

By (A.6), (A.10), and (A.11)-(A.13), $a_i^{BR}(\alpha_{-i}, \rho) \in \mathcal{A}_i$ for $i = 1, 2$.[21]

For any given arbitrary monotone strategies $\alpha \in \mathcal{A}$, denote the profile of best-response strategies by $a^{BR}(\alpha, \rho) \triangleq \{a_i^{BR}(\alpha_{-i}, \rho)\}_{i=1,2}$. Then, a BNE of $\mathcal{G}_\rho$, $a^\star(\rho)$, is given by the fixed point $a^{BR}(a^\star(\rho), \rho) = a^\star(\rho)$.

[21] By the monotonicity of the best response, a_i^{BR} is equivalent to the quantile function almost everywhere. We can thus directly use a_i^{BR} to characterize responsiveness by applying Lemma 1.

The proof to Theorem 2 proceeds in four steps:

1. Player i's best response strategy increases in responsiveness when player i's information quality increases (Lemma 3)

2. Player i's best response strategy increases in responsiveness when player $-i$ information quality increases (Lemma 4)

3. Player i's best response strategy increases in responsiveness when player $-i$'s strategy increases in responsiveness (Lemma 5)

4. Given 1-3, apply comparative statics on fixed points to get desired result.

We only prove the case for $u^i \in \Gamma^{\uparrow}$. A symmetric argument establishes the result for the case of $u^i \in \Gamma^{\downarrow}$.

Lemma 3 *Fix some arbitrary strategy $\alpha_{-i} \in \mathcal{A}_{-i}$. Consider two information structures $(\Sigma_{\rho_i'}, \Sigma_{\rho_{-i}})$ and $(\Sigma_{\rho_i''}, \Sigma_{\rho_{-i}})$ with $\rho_i'' \succeq_{MIO} \rho_i'$. If $u^i \in \Gamma^{\uparrow}$, then $a_i^{BR}(\alpha_{-i}, \rho_i'', \rho_{-i})$ is more responsive with a higher mean than $a_i^{BR}(\alpha_{-i}, \rho_i', \rho_{-i})$.*

Proof. Given $\Sigma_{\rho_{-i}}$ and $\alpha_{-i} \in \mathcal{A}_{-i}$, let

$$\tilde{u}^i(\theta_i, a_i) = \int_{\Theta_{-i} \times S_{-i}} u^i\Big(\theta_i, \alpha_{-i}(s_{-i}), a_i\Big) dF(s_{-i}|\theta_{-i}; \rho_{-i}) dF_{\Theta_{-i}}(\theta_{-i}|\theta_i)$$

and notice that

$$a_i^{BR}(s_i; \alpha_{-i}, \rho_i, \rho_{-i}) = \underset{a_i \in A_i}{\arg\max} \int_{\Theta_i} \tilde{u}^i(\theta_i, a_i) dF_{\Theta_i}(\theta_i|s_i; \rho_i).$$

We have mapped this problem to the single-agent framework where the payoff is given by $\tilde{u}^i : \Theta_i \times A_i \to \mathbb{R}$. Thus, if $\tilde{u}^i \in \mathcal{U}^{\uparrow}$, then $a_i^{BR}(\alpha_{-i}, \rho_i'', \rho_{-i})$ is more responsive with a higher mean than $a_i^{BR}(\alpha_{-i}, \rho_i', \rho_{-i})$ by Theorem 1.
First, notice that $\tilde{u}^i(\theta_i, a_i)$ inherits the measurability, boundedness, and smoothness properties of u^i. In particular $u^i_{a_i,a_i}(\theta_i, a_{-i}, a_i) < 0$ for all $(\theta_i, a_{-i}) \in \Theta_i \times A_{-i}$ implies that $\tilde{u}^i_{a_i,a_i}(\theta_i, a_i) < 0$ for all $\theta_i \in \Theta_i$. Similarly, $u^i_{a_i}(\theta_i, a_{-i}, a_i)$ is convex in a_i for all $(\theta_i, a_{-i}) \in \Theta_i \times A_{-i}$ implies that $\tilde{u}^i_{a_i}(\theta_i, a_i)$ is convex in a_i for all $\theta_i \in \Theta_i$.

To see supermodularity of $\tilde{u}^i$, let $\theta_i' > \theta_i$. Then,

$$\tilde{u}^i_{a_i}(\theta_i', a_i) - \tilde{u}^i_{a_i}(\theta_i, a_i)$$

$$= \int_{\Theta_{-i}\times S_{-i}} u^i_{a_i}\Big(\theta_i', \alpha_{-i}(s_{-i}), a_i\Big) dF(s_{-i}|\theta_{-i};\rho_{-i}) dF_{\Theta_{-i}}(\theta_{-i}|\theta_i')$$

$$- \int_{\Theta_{-i}\times S_{-i}} u^i_{a_i}\Big(\theta_i, \alpha_{-i}(s_{-i}), a_i\Big) dF(s_{-i}|\theta_{-i};\rho_{-i}) dF_{\Theta_{-i}}(\theta_{-i}|\theta_i)$$

$$= \int_{\Theta_{-i}\times S_{-i}} \Bigg(u^i_{a_i}\Big(\theta_i', \alpha_{-i}(s_{-i}), a_i\Big) - u^i_{a_i}\Big(\theta_i, \alpha_{-i}(s_{-i}), a_i\Big)\Bigg) dF(s_{-i}|\theta_{-i};\rho_{-i}) dF_{\Theta_{-i}}(\theta_{-i}|\theta_i')$$

$$+ \int_{\Theta_{-i}\times S_{-i}} u^i_{a_i}\Big(\theta_i, \alpha_{-i}(s_{-i}), a_i\Big) dF(s_{-i}|\theta_{-i};\rho_{-i}) \Bigg(dF_{\Theta_{-i}}(\theta_{-i}|\theta_i') - dF_{\Theta_{-i}}(\theta_{-i}|\theta_i)\Bigg)$$

Since $u^i(\theta_i, a_{-i}, a_i)$ has increasing differences (ID) in $(\theta_i; a_i)$ for each $a_{-i} \in A_{-i}$ and since ID is preserved under integration, the first term

$$\int_{\Theta_{-i}\times S_{-i}} \Bigg(u^i_{a_i}\Big(\theta_i', \alpha_{-i}(s_{-i}), a_i\Big) - u^i_{a_i}\Big(\theta_i, \alpha_{-i}(s_{-i}), a_i\Big)\Bigg) dF(s_{-i}|\theta_{-i};\rho_{-i}) dF_{\Theta_{-i}}(\theta_{-i}|\theta_i') \geq 0.$$

Furthermore, since $u^i(\theta_i, a_{-i}, a_i)$ has ID in $(a_{-i}; a_i)$ for each $\theta_i \in \Theta_i$, $u^i_{a_i}(\theta_i, a_{-i}, a_i)$ is increasing in a_{-i}. As α_{-i} is a monotone strategy, by (A.13) and (A.6), the second term

$$\int_{\Theta_{-i}\times S_{-i}} u^i_{a_i}\Big(\theta_i, \alpha_{-i}(s_{-i}), a_i\Big) dF(s_{-i}|\theta_{-i};\rho_{-i}) \Bigg(dF_{\Theta_{-i}}(\theta_{-i}|\theta_i') - dF_{\Theta_{-i}}(\theta_{-i}|\theta_i)\Bigg) \geq 0.$$

Hence, $\tilde{u}^i(\theta_i, a_i)$ is supermodular in (θ_i, a_i). A similar argument establishes that $\tilde{u}^i_{a_i}(\theta_i, a_i)$ is supermodular in (θ_i, a_i). Thus, $\tilde{u}^i \in \mathcal{U}^{\uparrow}$. The desired result in the statement of the lemma follows by Theorem 1. ■

Lemma 4 *Fix some arbitrary strategy $\alpha_{-i} \in \mathcal{A}_{-i}$. Consider two information structures $(\Sigma_{\rho_i}, \Sigma_{\rho'_{-i}})$ and $(\Sigma_{\rho_i}, \Sigma_{\rho''_{-i}})$ with $\rho''_{-i} \succeq_{MIO} \rho'_{-i}$. If $u^i \in \Gamma^{\uparrow}$, then $a_i^{BR}(\alpha_{-i}, \rho_i, \rho''_{-i})$ is more responsive with a higher mean than $a_i^{BR}(\alpha_{-i}, \rho_i, \rho'_{-i})$.*

Proof. Following the same first order condition argument we used in the proof of Theorem 1, we get the expression

$$\Big(a_i^{BR}(s_i;\alpha_{-i},\rho') - a_i^{BR}(s_i;\alpha_{-i},\rho'')\Big) \underbrace{\int_{\Theta\times S_{-i}} -u^i_{a_ia_i}\big(\theta_i,\alpha_{-i}(s_{-i}),a_i^{BR}(s_i;\alpha_{-i},\rho')\big)d\boldsymbol{F}(\theta,s_{-i}|s_i;\rho'')}_{\triangleq \hat{B}(s_i)}$$

$$+\int_{\Theta\times S_{-i}} u^i_{a_i}\big(\theta_i,\alpha_{-i}(s_{-i}),a_i^{BR}(s_i;\alpha_{-i},\rho')\big)\Big(d\boldsymbol{F}(\theta,s_{-i}|s_i;\rho'') - d\boldsymbol{F}(\theta,s_{-i}|s_i;\rho')\Big) \le 0.$$

Then, for each $t\in[0,1]$,

$$\begin{aligned}
&\int_t^1 \Big(a_i^{BR}(s_i;\alpha_{-i},\rho') - a_i^{BR}(s_i;\alpha_{-i},\rho'')\Big)ds_i \\
\le &\int_t^1 \hat{B}(s_i)^{-1}\int_{\Theta\times S_{-i}} u^i_{a_i}\big(\theta_i,\alpha_{-i}(s_{-i}),a_i^{BR}(s_i;\alpha_{-i},\rho')\big)\Big(d\boldsymbol{F}(\theta,s_{-i}|s_i;\rho') - d\boldsymbol{F}(\theta,s_{-i}|s_i;\rho'')\Big)ds_i \\
= &\int_{\Theta\times S} u^i_{a_i}\big(\theta_i,\alpha_{-i}(s_{-i}),a_i^{BR}(s_i;\alpha_{-i},\rho')\big)\hat{B}(s_i)^{-1}\mathbb{1}_{[s_i\ge t]}\Big(d\boldsymbol{F}(\theta,s;\rho') - d\boldsymbol{F}(\theta,s;\rho'')\Big).
\end{aligned}$$

For a given information structure Σ_ρ,

$$\boldsymbol{F}(\theta,s;\rho) = \int_{\underline{\theta}_{-i}}^{\theta_{-i}}\int_0^{s_{-i}} \boldsymbol{F}(\theta_i,s_i|\omega_{-i};\rho)dF(\omega_{-i},x_{-i};\rho_{-i})$$

where the equality follows from (A.11). Let

$$\hat{\psi}(\theta_{-i},s_{-i};t) = \int_{\Theta_i\times S_i} u^i_{a_i}\big(\theta_i,\alpha_{-i}(s_{-i}),a_i^{BR}(s_i;\alpha_{-i},\rho')\big)\hat{B}(s_i)^{-1}\mathbb{1}_{[s_i\ge t]}d\boldsymbol{F}(\theta_i,s_i|\theta_{-i};\rho)$$

so that

$$\begin{aligned}
&\int_t^1 \Big(a_i^{BR}(s_i;\alpha_{-i},\rho') - a_i^{BR}(s_i;\alpha_{-i},\rho'')\Big)ds_i \\
\le &\int_{\Theta_{-i}\times S_{-i}} \hat{\psi}(\theta_{-i},s_{-i};t)\Big(dF(\theta_{-i},s_{-i};\rho'_{-i}) - dF(\theta_{-i},s_{-i};\rho''_{-i})\Big).
\end{aligned}$$

Take $s'_{-i} > s_{-i}$ which implies that $\alpha_{-i}(s'_{-i}) \geq \alpha_{-i}(s_{-i})$. Then,

$$\left[u^i_{a_i}\big(\theta_i, \alpha_{-i}(s'_{-i}), a_i^{BR}(s_i;\alpha_{-i},\rho')\big) - u^i_{a_i}\big(\theta_i, \alpha_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha_{-i},\rho')\big)\right]\mathbb{1}_{[s_i \geq t]} \geq 0,$$

for all $(\theta_i, s_i) \in \Theta_i \times S_i$ because u^i has increasing differences in $(a_i; a_{-i})$. It is also increasing in both θ_i and s_i because $u^i_{a_i}$ has increasing differences in $(\theta_i; a_{-i})$ and $(a_i; a_{-i})$. Similarly,

$$\hat{B}(s_i)^{-1} \geq 0$$

by concavity of u^i in a_i and it is increasing in s_i because $u^i_{a_i}$ has ID in $(\theta_i; a_i)$ and $(a_i; a_{-i})$, and because $\boldsymbol{F}(\theta, s_{-i}|s_i)$ is increasing in FOSD as s_i increases. Thus, along with (A.6) and (A.13),

$$\begin{aligned}
&\hat{\psi}(\theta_{-i}, s'_{-i}; t) - \hat{\psi}(\theta_{-i}, s_{-i}; t) \\
&= \int_{\Theta_i \times S_i} \left[u^i_{a_i}\big(\theta_i, \alpha_{-i}(s'_{-i}), a_i^{BR}(s_i;\alpha_{-i},\rho')\big) - u^i_{a_i}\big(\theta_i, \alpha_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha_{-i},\rho')\big)\right] \\
&\qquad \times \hat{B}(s_i)^{-1}\mathbb{1}_{[s_i \geq t]} d\boldsymbol{F}(\theta_i, s_i|\theta_{-i};\rho)
\end{aligned}$$

is increasing in θ_{-i}. In other words, $\hat{\psi}(\theta_{-i}, s_{-i}; t)$ is supermodular in (θ_{-i}, s_{-i}). By Lemma 2, $\rho''_{-i} \succeq_{MIO} \rho'_{-i}$ implies

$$\int_{\Theta_{-i} \times S_{-i}} \hat{\psi}(\theta_{-i}, s_{-i}; t)\Big(dF(\theta_{-i}, s_{-i}; \rho'_{-i}) - dF(\theta_{-i}, s_{-i}; \rho''_{-i})\Big) \leq 0,$$

giving us the desired result. ■

Lemma 5 *Fix Σ_ρ. Let $\alpha''_{-i}, \alpha'_{-i} \in \mathcal{A}_{-i}$ such that α''_{-i} is more responsive with higher mean than α'_{-i}. If $u^i \in \Gamma^{\uparrow}$, then, $a_i^{BR}(\alpha''_{-i}, \rho)$ is more responsive with a higher mean than $a_i^{BR}(\alpha'_{-i}, \rho)$.*

Proof. Suppress the dependence on ρ as it is held fixed. For any $t \in [0,1]$, we use the first order conditions argument (similar to the proof of Lemma 4) to get the expression

$$
\begin{aligned}
&\int_t^1 \left(a_i^{BR}(s_i;\alpha'_{-i}) - a_i^{BR}(s_i;\alpha''_{-i})\right)ds_i \\
\leq &\int_t^1 \Bigg\{ \Bigg(\underbrace{-\int_{\Theta\times S_{-i}} u^i_{a_i a_i}\Big(\theta, \alpha''_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\Big) d\boldsymbol{F}(\theta, s_{-i}|s_i)}_{\triangleq \tilde{B}_i(s_i)} \Bigg)^{-1} \\
\times &\int_{\Theta\times S_{-i}} \left[u^i_{a_i}\Big(\theta_i, \alpha'_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\Big) - u^i_{a_i}\Big(\theta_i, \alpha''_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\Big) \right] d\boldsymbol{F}(\theta, s_{-i}|s_i) \Bigg\} ds_i.
\end{aligned}
$$

Since $u^i_{a_i}$ is continuous and increasing in a_{-i} (by ID of u^i in $(a_{-i};a_i)$), it is differentiable in a_{-i} almost everywhere. By convexity of $u^i_{a_i}$ in a_{-i},

$$
\begin{aligned}
&u^i_{a_i}\Big(\theta_i, \alpha'_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\Big) - u^i_{a_i}\Big(\theta_i, \alpha''_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\Big) \\
\leq &u^i_{a_i a_{-i}}\Big(\theta_i, \alpha'_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\Big)\big(\alpha'_{-i}(s_{-i}) - \alpha''_{-i}(s_{-i})\big).
\end{aligned}
$$

Thus,

$$
\begin{aligned}
&\int_t^1 \left(a_i^{BR}(s_i;\alpha'_{-i}) - a_i^{BR}(s_i;\alpha''_{-i})\right)ds_i \\
\leq &\int_{S_{-i}} \left(\alpha'_{-i}(s_{-i}) - \alpha''_{-i}(s_{-i})\right) \int_{\Theta\times S_i} u^i_{a_i a_{-i}}\Big(\theta_i, \alpha'_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\Big)\tilde{B}(s_i)^{-1}\mathbb{1}_{[s_i\geq t]} d\boldsymbol{F}(\theta, s_i|s_{-i}) ds_{-i}.
\end{aligned}
$$

We make use of the following result from Quah and Strulovici (2009)

Lemma 6 *Let $g:[x',x'']\to\mathbb{R}$ and $h:[x',x'']\to\mathbb{R}$ be integrable functions.*

1. *If g is increasing and $\int_x^{x''} h(t)dt \geq 0$ for all $x\in[x',x'']$, then $\int_{x'}^{x''} g(t)h(t)dt \geq g(x')\int_{x'}^{x''} h(t)dt$*
2. *If g is decreasing and $\int_{x'}^{x} h(t)dt \geq 0$ for all $x\in[x',x'']$, then $\int_{x'}^{x''} g(t)h(t)dt \geq g(x'')\int_{x'}^{x''} h(t)dt$*

Proof. Quah and Strulovici (2009) Lemma 1 ■

By using the definition of responsiveness in Lemma 1 and the equivalence of the monotone strategy α_{-i} with its quantile function, α''_{-i} is more responsive with a higher mean than α'_{-i} if, and only if,

$$\int_t^1 \left(\alpha'_{-i}(s_{-i}) - \alpha''_{-i}(s_{-i})\right)ds_{-i} \leq 0,\ \ \forall t\in[0,1].$$

Furthermore,

$$u^i_{a_i a_{-i}}\Big(\theta_i, \alpha'_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\Big)\mathbb{1}_{[s_i \geq t]} \geq 0,\ \forall(\theta_i, s_i) \in \Theta_i \times S_i$$

as u^i has increasing differences in $(a_i; a_{-i})$ for all $\theta_i \in \Theta_i$. It is also increasing in both θ_i and s_i because $u^i_{a_i}$ has increasing differences in $(\theta_i; a_{-i})$ and $(a_i; a_{-i})$. Similarly,

$$\tilde{B}(s_i)^{-1} \geq 0$$

by concavity of u^i in a_i and it is increasing in s_i because $u^i_{a_i}$ has ID in $(\theta_i; a_i)$ and $(a_i; a_{-i})$, and because $\boldsymbol{F}(\theta, s_{-i}|s_i)$ is increasing in FOSD as s_i increases. Thus, along with (A.6) and (A.13),

$$\int_{\Theta\times S_i} u^i_{a_i a_{-i}}\Big(\theta_i, \alpha'_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\Big)\tilde{B}(s_i)^{-1}\mathbb{1}_{[s_i\geq t]}d\boldsymbol{F}(\theta, s_i|s_{-i})$$

is an increasing function of s_{-i}. Applying Lemma 6, we have

$$\begin{aligned}
&\int_t^1 \big(a_i^{BR}(s_i;\alpha'_{-i}) - a_i^{BR}(s_i;\alpha''_{-i})\big)ds_i \\
\leq &\int_{S_{-i}} \big(\alpha'_{-i}(s_{-i}) - \alpha''_{-i}(s_{-i})\big)\int_{\Theta\times S_i} u^i_{a_i a_{-i}}\Big(\theta_i, \alpha'_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\Big)\tilde{B}(s_i)^{-1}\mathbb{1}_{[s_i\geq t]}d\boldsymbol{F}(\theta, s_i|s_{-i})ds_{-i} \\
\leq &\underbrace{\int_{S_{-i}} \big(\alpha'_{-i}(s_{-i}) - \alpha''_{-i}(s_{-i})\big)ds_{-i}}_{\leq 0}\int_{\Theta\times S_i} \underbrace{u^i_{a_i a_{-i}}\Big(\theta_i, \alpha'_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\Big)\tilde{B}(s_i)^{-1}\mathbb{1}_{[s_i\geq t]}}_{\geq 0}d\boldsymbol{F}(\theta, s_i|0) \\
\leq &0
\end{aligned}$$

for each $t \in [0, 1]$. ■

We will now tackle the last step in the proof: comparative statics of the BNEs. We apply the comparative statics of fixed points provided by Villas-Boas (1997). To do so, we will need the following definition.

Definition 1 (Contractible Space) *Let X be a topological space. We say that X is a contractible space if there exists a map $\Phi : X \times [0, 1] \to X$ such that for all $x \in X$*

1. $\Phi(\cdot,\lambda)$ *is continuous in* λ*,*

2. $\Phi(x,0) = x$ *and* $\Phi(x,1) = x^*$ *for some* $x^* \in X$

Intuitively, X is contractible if it can be continuously shrunk into a point inside itself.

Theorem 6, Villas-Boas (1997): *Let X be a compact subset of a Banach space. Consider continuous mappings $T_1 : X \to X$ and $T_2 : X \to X$, and a transitive and reflexive order $\succeq$ on X. For all $x \in X$, let the upper-set $\{x' \in X : x' \succeq x\}$ be a compact and contractible subset. Let both T_1 and T_2 have a fixed point on X. Suppose $x' \succeq x \Rightarrow T_1(x') \succeq T_1(x)$, and suppose $T_1(x) \succeq T_2(x)$ for all $x \in X$. Then for every fixed point $x_2^\star$ of T_2, there is a fixed point $x_1^\star$ of T_1 such that $x_1^\star \succeq x_2^\star$.*

The remaining few steps prove that our setting satisfies the assumptions needed to apply the Villas-Boas result.[22] Let $BV([0,1])$ be the space of functions of bounded variation from $[0,1]$ to $\mathbb{R}$. Given a function $g \in BV([0,1])$, let $V(g)$ be the total variation of g.[23] Define the bounded variation norm by $||g||_{BV} = \int_0^1 |g(s)|ds + V(g)$. The space $BV([0,1])$ equipped with the $||\cdot||_{BV}$ is a Banach space.

Fix a player $i \in N$. Any $\alpha_i \in \mathcal{A}_i$ is both uniformly bounded and is of bounded variation as it is an increasing function with $V(\alpha_i) = \alpha_i(1) - \alpha_i(0) \leq \bar{a}_i - \underline{a}_i < \infty$. Therefore, $\mathcal{A}_i$ is a subset of $BV([0,1])$.

Lemma 7 *$\mathcal{A}_i$ is a compact subset of $BV([0,1])$.*

Proof. We first show that $\mathcal{A}_i$ is a closed subset $BV([0,1])$. Take a sequence $\{\tilde{\alpha}_{i,k}\}_{k=1}^\infty \in \mathcal{A}_i$ that converges to $\tilde{\alpha}_i \in BV([0,1])$ in the $||\cdot||_{BV}$ norm. Thus, $\tilde{\alpha}_{i,k}$ converges point-wise to $\tilde{\alpha}_i$ almost everywhere. The point-wise limit of monotone functions is also monotone. Furthermore, as $\underline{a}_i \leq \tilde{\alpha}_{i,k}(0)$ for all k, the limit also satisfies $\underline{a}_i \leq \tilde{\alpha}_i(0)$. Similarly, as $\bar{a}_i \geq \tilde{\alpha}_{i,k}(1)$ for all k, the limit also satisfies $\bar{a}_i \geq \tilde{\alpha}_i(1)$. Finally, the point-wise limit of measurable functions is measurable (Corollary 8.9, Measure, Integrals, and Martingales, Schilling, 2005). As $\tilde{\alpha}_i$ is a monotone and measurable function that maps from $[0,1]$ to A_i, $\tilde{a}_i \in \mathcal{A}_i$.

[22]For the case when $u^i \in \Gamma^\downarrow$ for all $i \in N$, we use Theorem 7 of Villas-Boas (1997) which uses the lower-sets generated by responsiveness with a lower mean to get the desired comparative statics of fixed points.

[23]Specifically, $V(g) = \sup_{p \in P} \sum_{i=0}^{n_p-1} |g(x_{i+1}) - g(x_i)|$ where P is the set of all partitions $p = \{x_0, x_1, \ldots, x_{n_p}\}$ on $[0,1]$.

Now, consider any sequence $\{\alpha_{i,k}\}_{k=1}^{\infty} \in \mathcal{A}_i \subset BV([0,1])$. Using, Helly's Selection Theorem, there exists a subsequence $\{\alpha_{i,k_m}\}_{k_m} \in \mathcal{A}_i$ that converges point-wise to a limit $\alpha_i \in BV([0,1])$. Notice that

$$||\alpha_{i,k_m} - \alpha_i||_{BV} = \int_0^1 \big|\alpha_{i,k_m}(s) - \alpha_i(s)\big| ds + V(\alpha_{i,k_m} - \alpha_i)$$

$$\leq \int_0^1 \big|\alpha_{i,k_m}(s) - \alpha_i(s)\big| ds + V(\alpha_{i,k_m}) - V(\alpha_i)$$

$$= \int_0^1 \big|\alpha_{i,k_m}(s) - \alpha_i(s)\big| ds + \Big(\alpha_{i,k_m}(1) - \alpha_{i,k_m}(0)\Big) - \Big(\alpha_i(1) - \alpha_i(0)\Big)$$

where the inequality follows from the sub-additive property of the total variation function, i.e., $V(g_1 + g_2) \leq V(g_1) + V(g_2)$, and the equality follows as both α_{i,k_m} and the point-wise limit α_i are increasing functions. By the dominated convergence theorem, $\int_0^1 \big|\alpha_{i,k_m}(s) - \alpha_i(s)\big| ds$ converges to zero as k_m goes to infinity. Similarly, $\alpha_{i,k_m}(1) - \alpha_i(1)$ and $\alpha_{i,k_m}(0) - \alpha_i(0)$ converge to zero as k_m goes to infinity. Hence, α_{i,k_m} converges to α_i in $|| \cdot ||_{BV}$. As $\mathcal{A}_i$ is closed, $\alpha_i \in \mathcal{A}_i$. Therefore, $\mathcal{A}_i$ is (sequentially) compact.

■

Define a partial order over $\mathcal{A}_i$ by $\alpha_i' \succeq_i \alpha_i$ if, and only if, α_i' is more responsive with a higher mean than α_i. Denote the upper-set of a_i by $\mathcal{US}(\alpha_i) = \{\alpha_i' \in \mathcal{A}_i : \alpha_i' \succeq_i \alpha_i\} \subseteq \mathcal{A}_i$.

Lemma 8 *For any $\alpha_i \in \mathcal{A}_i$, the upper-set $\mathcal{US}(\alpha_i)$ is a compact and contractible set.*

Proof. For some $\alpha_i \in \mathcal{A}_i$, the upper-set, $\mathcal{US}(\alpha_i)$, is a closed subset of $\mathcal{A}_i$ (follows from the dominated convergence Theorem). Hence, it is compact. To show that $\mathcal{US}(\alpha_i)$ is contractible, let $\alpha_i^c : [0,1] \to A_i$ be the constant function with $\alpha_i^c(s) = \bar{a}_i$ for all $s \in [0,1]$. Note that $\alpha_i^c \in \mathcal{A}_i$. Furthermore, $\alpha_i^c(s) \geq \alpha_i(s)$, $\forall s \in [0,1]$ which implies $\alpha_i^c \succeq_i \alpha_i \Rightarrow \alpha_i^c \in \mathcal{US}(\alpha_i)$.
For each $\alpha_i \in \mathcal{A}_i$, defne the mapping $\Phi : \mathcal{US}(\alpha_i) \times [0,1] \to \mathcal{US}(\alpha_i)$ such that

$$\Phi(\alpha_i', \lambda) = (1-\lambda)\alpha_i' + \lambda\alpha_i^c.$$

$\Phi(\cdot, \lambda)$ is continuous in λ. As λ increases from 0 to 1, Φ continuously deforms any strategy in $\mathcal{US}(\alpha_i)$ to the constant strategy α_i^c, which is itself in $\mathcal{US}(\alpha_i)$. Therefore, $\mathcal{US}(\alpha_i)$ is contractible.

■

We therefore have an order, $\succsim_i$, on $\mathcal{A}_i$ (which is a subset of a Banach space) that generates compact and contractible upper-sets. We extend these properties to $\mathcal{A} = \times_{i=1}^{I} \mathcal{A}_i$ by the product order: given $\alpha, \alpha' \in \mathcal{A}$, $\alpha' \succeq \alpha$ if, and only if, $\alpha'_i \succeq_i \alpha_i$ for all $i \in N$. Along with the product topology, $\succeq$ is an order on $\mathcal{A}$ that generates compact and contractible upper-sets.[24]
Consider a Bayesian game $\mathcal{G}_\rho = (\Sigma_\rho, G)$. Define an operator $T_\rho : \mathcal{A} \to \mathcal{A}$ with

$$T_\rho(\alpha) = \big(a_1^{BR}(\alpha_{-1}, \rho), a_2^{BR}(\alpha_{-2}, \rho), \ldots, a_n^{BR}(\alpha_{-n}, \rho)\big).$$

T_ρ is continuous in α as utility functions are continuous in actions. A monotone BNE of $\mathcal{G}_\rho$ is a fixed point of T_ρ. We know such a fixed point exists (Van Zandt and Vives (2007).
Now consider two different games, $\mathcal{G}_{\rho''} = (\Sigma_{\rho''}, G)$ and $\mathcal{G}_{\rho'} = (\Sigma_{\rho'}, G)$, with $\rho'' \succeq_{MIO} \rho'$. From Lemma 5,

$$\alpha' \succeq \alpha \Leftrightarrow \alpha'_i \succeq_i \alpha_i, \forall i \in N \Rightarrow a_i^{BR}(\alpha'_{-i}, \rho'') \succeq_i a_i^{BR}(\alpha_{-i}, \rho''), \forall i \in N \Leftrightarrow T_{\rho''}(\alpha') \succeq T_{\rho''}(\alpha).$$

From Lemma 3 and 4,

$$\rho'' \succeq_{MIO} \rho' \Rightarrow a_i^{BR}(\alpha_{-i}, \rho'') \succeq_i a_i^{BR}(\alpha_{-i}, \rho'), \forall i \in N \Leftrightarrow T_{\rho''}(\alpha) \succeq T_{\rho'}(\alpha)$$

for all $\alpha \in \mathcal{A}$. We can now directly apply Theorem 6 of Villas-Boas (1997) to conclude that, for every fixed point $a^\star(\rho')$ of $T_{\rho'}$, there is a fixed point $a^\star(\rho'')$ of $T_{\rho''}$ such that $a^\star(\rho'') \succeq a^\star(\rho')$.
■

Proof of Proposition 3

[24] $\mathcal{A}$ is a subset of a Banach space equipped with the metric $d(\alpha', \alpha) = \sum_i ||\alpha'_i - \alpha_i||_{BV}$.

Proof. Take any two information structures $\Sigma_{\rho'}$, $\Sigma_{\rho''} \in \mathcal{P}$ with $\rho'' \succeq_{MIO} \rho'$. Then,

$$U_1(\rho'') - U_1(\rho')$$

$$= \int_\Theta \int_S u^1\big(\theta, a_1^\star(\theta;\rho''), a_2^\star(s;\rho'')\big) dF(s|\theta;\rho'')\mu^o(d\theta) - \int_\Theta \int_S u^1\big(\theta, a_1^\star(\theta;\rho'), a_2^\star(s;\rho')\big) dF(s|\theta;\rho')\mu^o(d\theta)$$

$$= \int_\Theta \int_S \Big(u^1\big(\theta, a_1^\star(\theta;\rho''), a_2^\star(s;\rho'')\big) - u^1\big(\theta, a_1^\star(\theta;\rho'), a_2^\star(s;\rho'')\big)\Big) dF(s|\theta;\rho'')\mu^o(d\theta)$$

$$+ \int_\Theta \int_S \Big(u^1\big(\theta, a_1^\star(\theta;\rho'), a_2^\star(s;\rho'')\big) - u^1\big(\theta, a_1^\star(\theta;\rho'), a_2^\star(s;\rho')\big)\Big) dF(s|\theta;\rho'')\mu^o(d\theta)$$

$$+ \int_\Theta \int_S u^1\big(\theta, a_1^\star(\theta;\rho'), a_2^\star(s;\rho')\big)\Big) \Big(dF(s|\theta;\rho'') - dF(s|\theta;\rho')\Big)\mu^o(d\theta).$$

The first term is non-negative as $a_1^\star(\rho'')$ is player 1's best response to $a_2^\star(\rho'')$ and information structure $\Sigma_{\rho''}$. After changing the order of integration, the second term can be rewritten as

$$\int_S \int_\Theta v^1(\theta) a_1^\star(\theta;\rho')\mu(d\theta|s;\rho)\Big(a_2^\star(s;\rho'') - a_2^\star(s;\rho')\Big)ds.$$

By (A.12),

$$\int_\Theta v^1(\theta) a_1^\star(\theta;\rho')\mu(d\theta|s;\rho)$$

is increasing in s. Furthermore, because $u^i \in \Gamma^\uparrow$ for $i = 1, 2$, by Theorem 2, $\rho'' \succeq_{MIO} \rho'$ implies $a_2^\star(\rho'')$ is more responsive with a higher mean than $a_2^\star(\rho')$. By Lemma 1,

$$\int_t^1 \Big(a_2^\star(s;\rho'') - a_2^\star(s;\rho')\Big)ds \geq 0$$

for all $t \in [0, 1]$. By Lemma 6 (Lemma 1 from Quah and Strulovici (2009)), the second term is then

$$\int_S \int_\Theta v^1(\theta) a_1^\star(\theta;\rho')\mu(d\theta|s;\rho)\Big(a_2^\star(s;\rho'') - a_2^\star(s;\rho')\Big)ds$$

$$\geq \int_\Theta v^1(\theta) a_1^\star(\theta;\rho')\mu(d\theta|0;\rho)\int_S \Big(a_2^\star(s;\rho'') - a_2^\star(s;\rho')\Big)ds \geq 0.$$

For the third term, take $s'' > s'$ and note that

$$u^1\big(\theta, a_1^\star(\theta;\rho'), a_2^\star(s'';\rho')\big) - u^1\big(\theta, a_1^\star(\theta;\rho'), a_2^\star(s';\rho')\big)$$

$$= v^1(\theta) a_1^\star(\theta;\rho') \underbrace{\Big(a_2^\star(s'';\rho') - a_2^\star(s';\rho')\Big)}_{\geq 0}$$

is increasing in θ. Hence, $u^1\big(\theta, a_1^\star(\theta;\rho'), a_2^\star(s;\rho')\big)$ is supermodular in (θ, s). By Lemma 2,

$$\int_{\Theta\times S} u^1\big(\theta, a_1^\star(\theta;\rho'), a_2^\star(s;\rho')\big)\Big(dF(s|\theta;\rho'') - dF(s|\theta;\rho')\Big)\mu^o(d\theta) \geq 0.$$

Thus, $\rho'' \succeq_{MIO} \rho'$ implies $U_1(\rho'') \geq U_1(\rho')$. As $\bar{\rho} \succeq_{MIO} \rho$ for all $\rho \in \mathcal{P}$, player 1's ex-ante payoff is maximized by the full-information structure. ■

6.3 Proofs from Section 4

Proof of Proposition 4

Proof. Notice that we can rewrite $U_1(\rho;\rho) - U_1(\hat{\rho};\hat{\rho})$ as

$$\underbrace{U_1(\rho;\hat{\rho}) - U_1(\hat{\rho};\hat{\rho})}_{\text{value of covert information}} + \underbrace{U_1(\rho;\rho) - U_1(\rho;\hat{\rho})}_{=VT(\rho;\hat{\rho})}.$$

Amir and Lazzati, (2016, Proposition 7) show that the first term is non-negative when $\rho_1 \succeq_{MIO} \hat{\rho}_1$, i.e., the value of covert information is non-negative when quality of information increases. Hence, if $VT(\rho;\hat{\rho}) \geq 0$, we can conclude that the value of overt information is also non-negative when quality of information increases. ■

Proof of Proposition 5

Proof. Suppose $\Sigma_{\rho_1^c} \neq \Sigma_{\rho_1^o}$ (otherwise, it is trivial). By definition,

$$U_1(\rho^c;\rho^c) - \kappa(\rho_1^c) \geq U_1(\rho^o;\rho^c) - \kappa(\rho_1^o)$$

and

$$U_1(\rho^o;\rho^o) - \kappa(\rho_1^o) \geq U_1(\rho^c;\rho^c) - \kappa(\rho_1^c).$$

Combining the two inequalities, we get

$$U_1(\rho^o;\rho^o) - U_1(\rho^o;\rho^c) = VT(\rho^o;\rho^c) \geq 0 \Leftrightarrow \rho_1^o \succeq_{MIO} \rho_1^c.$$

■

Proof of Theorem 3

Proof. We only show the proof for the case when $u^1(\theta_1, a)$ is an increasing and convex function of a_2 and $a_2^*(\cdot)$ becomes more responsive with a higher mean as information quality increases. A similar argument applies when $u^1(\theta_1, a)$ is a decreasing concave function of a_2, and $a_2^*(\cdot)$ becomes more responsive with a lower mean as information quality increases.

($\Longrightarrow$) Suppose $\rho_1 \succeq_{MIO} \hat{\rho}_1$ and $a_2^\star(\rho)$ is more responsive with a higher mean than $a_2^\star(\hat{\rho})$. $VT(\rho;\hat{\rho}) = U_1(\rho;\rho) - U_1(\rho;\hat{\rho})$ is given by

$$\int_{\Theta\times S} \Big[u^1\big(\theta_1, a_1^\star(s_1;\rho), a_2^\star(s_2;\rho)\big) - u^1\big(\theta_1, a_1^{BR}(s_1; a_2^\star(\hat{\rho}), \rho), a_2^\star(s_2;\hat{\rho})\big)\Big] d\boldsymbol{F}(\theta, s;\rho)$$

$$= \underbrace{\int_{\Theta\times S} \Big[u^1\big(\theta_1, a_1^\star(s_1;\rho), a_2^\star(s_2;\rho)\big) - u^1\big(\theta_1, a_1^{BR}(s_1; a_2^\star(\hat{\rho}), \rho), a_2^\star(s_2;\rho)\big)\Big] d\boldsymbol{F}(\theta, s;\rho)}_{\geq 0 \text{ by optimality}}$$

$$+ \int_{\Theta\times S} \Big[u^1\big(\theta_1, a_1^{BR}(s_1; a_2^\star(\hat{\rho}), \rho), a_2^\star(s_2;\rho)\big) - u^1\big(\theta_1, a_1^{BR}(s_1; a_2^\star(\hat{\rho}), \rho), a_2^\star(s_2;\hat{\rho})\big)\Big] d\boldsymbol{F}(\theta, s;\rho)$$

$$\geq \int_{\Theta\times S} \Big[u^1\big(\theta_1, a_1^{BR}(s_1; a_2^\star(\hat{\rho}), \rho), a_2^\star(s_2;\rho)\big) - u^1\big(\theta_1, a_1^{BR}(s_1; a_2^\star(\hat{\rho}), \rho), a_2^\star(s_2;\hat{\rho})\big)\Big] d\boldsymbol{F}(\theta, s;\rho).$$

Suppose u^1 is an increasing convex function of a_2. Note that $u^1_{a_2} \geq 0$ implies that player 2's actions have positive externalities on player 1's payoff. By convexity,

$$\int_{\Theta\times S}\Big[u^1\big(\theta_1,a_1^{BR}(s_1;a_2^\star(\hat\rho),\rho),a_2^\star(s_2;\rho)\big)-u^1\big(\theta_1,a_1^{BR}(s_1;a_2^\star(\hat\rho),\rho),a_2^\star(s_2;\hat\rho)\big)\Big]d\boldsymbol{F}(\theta,s;\rho)$$

$$\geq\int_{S_2}\Big(a_2^\star(s_2;\rho)-a_2^\star(s_2;\hat\rho)\Big)\int_{\Theta_1\times S_1}u^1_{a_2}\big(\theta_1,a_1^{BR}(s_1;a_2^\star(\hat\rho),\rho),a_2^\star(s_2;\hat\rho)\big)d\boldsymbol{F}(\theta_1,s_1|s_2;\rho)ds_2.$$

If states are independently distributed across players, given (A.11),

$$\int_{S_2}\Big(a_2^\star(s_2;\rho)-a_2^\star(s_2;\hat\rho)\Big)\int_{\Theta_1\times S_1}u^1_{a_2}\big(\theta_1,a_1^{BR}(s_1;a_2^\star(\hat\rho),\rho),a_2^\star(s_2;\hat\rho)\big)d\boldsymbol{F}(\theta_1,s_1|s_2;\rho)ds_2$$

$$=\underbrace{\int_{S_2}\Big(a_2^\star(s_2;\rho)-a_2^\star(s_2;\hat\rho)\Big)ds_2}_{\substack{\geq 0\text{ by responsiveness}\\ \text{with a higher mean}}}\int_{\Theta_1\times S_1}\underbrace{u^1_{a_2}\big(\theta_1,a_1^{BR}(s_1;a_2^\star(\hat\rho),\rho),a_2^\star(s_2;\hat\rho)\big)}_{\geq 0\text{ by positive externalities}}dF(\theta_1,s_1;\rho_1)\geq 0.$$

If $u^1(\theta_1,a)$ has ID in $(\theta_1;a_2)$, then

$$\int_{\Theta_1\times S_1}u^1_{a_2}\big(\theta_1,a_1^{BR}(s_1;a_2^\star(\hat\rho),\rho),a_2^\star(s_2;\hat\rho)\big)d\boldsymbol{F}(\theta_1,s_1|s_2;\rho)$$

in an increasing function of s_2. From Lemma 1, responsiveness with a higher mean is equivalent to

$$\int_t^1\Big(a_2^\star(s_2;\rho)-a_2^\star(s_2;\hat\rho)\Big)ds_2\geq 0$$

for all $t\in[0,1]$. Using Lemma 6, we can then conclude that

$$\int_{S_2}\Big(a_2^\star(s_2;\rho)-a_2^\star(s_2;\hat\rho)\Big)\int_{\Theta_1\times S_1}u^1_{a_2}\big(\theta_1,a_1^{BR}(s_1;a_2^\star(\hat\rho),\rho),a_2^\star(s_2;\hat\rho)\big)d\boldsymbol{F}(\theta_1,s_1|s_2;\rho)ds_2$$

$$\geq\underbrace{\int_{S_2}\Big(a_2^\star(s_2;\rho)-a_2^\star(s_2;\hat\rho)\Big)ds_2}_{\substack{\geq 0\text{ by responsiveness}\\ \text{with a higher mean}}}\int_{\Theta_1\times S_1}\underbrace{u^1_{a_2}\big(\theta_1,a_1^{BR}(s_1;a_2^\star(\hat\rho),\rho),a_2^\star(0;\hat\rho)\big)}_{\geq 0\text{ by positive externalities}}d\boldsymbol{F}(\theta_1,s_1|0;\rho)\geq 0.$$

In either case, we get the desired result that $VT(\rho;\hat\rho)\geq 0$.
($\Longleftarrow$) Suppose $\rho_1\not\succeq_{MIO}\hat\rho_1\implies\hat\rho_1\succeq_{MIO}\rho_1$ because $\mathcal{P}$ is a totally ordered set of information

structures. By assumption, $a_2^\star(\hat\rho)$ is more responsive with a higher mean than $a_2^\star(\rho)$. Then, $-VT(\rho;\hat\rho) = U_1(\rho;\hat\rho) - U_1(\rho;\rho)$ is given by

$$\int_{\Theta\times S}\Big[u^1\big(\theta_1, a_1^{BR}(s_1;a_2^\star(\hat\rho),\rho), a_2^\star(s_2;\hat\rho)\big) - u^1\big(\theta_1, a_1^\star(s_1;\rho), a_2^\star(s_2;\rho)\big)\Big]d\boldsymbol{F}(\theta,s;\rho)$$

$$=\underbrace{\int_{\Theta\times S}\Big[u^1\big(\theta_1, a_1^{BR}(s_1;a_2^\star(\hat\rho),\rho), a_2^\star(s_2;\hat\rho)\big) - u^1\big(\theta_1, a_1^\star(s_1;\rho), a_2^\star(s_2;\hat\rho)\big)\Big]d\boldsymbol{F}(\theta,s;\rho)}_{\geq 0 \text{ by optimality}}$$

$$+\int_{\Theta\times S}\Big[u^1\big(\theta_1, a_1^\star(s_1;\rho), a_2^\star(s_2;\hat\rho)\big) - u^1\big(\theta_1, a_1^\star(s_1;\rho), a_2^\star(s_2;\rho)\big)\Big]d\boldsymbol{F}(\theta,s;\rho)$$

$$\geq\int_{\Theta\times S}\Big[u^1\big(\theta_1, a_1^\star(s_1;\rho), a_2^\star(s_2;\hat\rho)\big) - u^1\big(\theta_1, a_1^\star(s_1;\rho), a_2^\star(s_2;\rho)\big)\Big]d\boldsymbol{F}(\theta,s;\rho).$$

Suppose u^1 is an increasing convex function of a_2. By convexity,

$$\int_{\Theta\times S}\Big[u^1\big(\theta_1, a_1^\star(s_1;\rho), a_2^\star(s_2;\hat\rho)\big) - u^1\big(\theta_1, a_1^\star(s_1;\rho), a_2^\star(s_2;\rho)\big)\Big]d\boldsymbol{F}(\theta,s;\rho)$$

$$\geq\int_{S_2}\Big(a_2^\star(s_2;\hat\rho) - a_2^\star(s_2;\rho)\Big)\int_{\Theta_1\times S_1} u^1_{a_2}\big(\theta_1, a_1^\star(s_1;\rho), a_2^\star(s_2;\rho)\big)d\boldsymbol{F}(\theta_1,s_1|s_2;\rho)ds_2.$$

If states are independently distributed across players, then by (A.11),

$$\int_{S_2}\Big(a_2^\star(s_2;\hat\rho) - a_2^\star(s_2;\rho)\Big)\int_{\Theta_1\times S_1} u^1_{a_2}\big(\theta_1, a_1^\star(s_1;\rho), a_2^\star(s_2;\rho)\big)d\boldsymbol{F}(\theta_1,s_1|s_2;\rho)ds_2$$

$$=\underbrace{\int_{S_2}\Big(a_2^\star(s_2;\hat\rho) - a_2^\star(s_2;\rho)\Big)ds_2}_{\substack{\geq 0 \text{ by responsiveness}\\ \text{with a higher mean}}}\int_{\Theta_1\times S_1}\underbrace{u^1_{a_2}\big(\theta_1, a_1^\star(s_1;\rho), a_2^\star(s_2;\rho)\big)}_{\geq 0 \text{ by positive externalities}}dF(\theta_1,s_1;\rho_1)\geq 0.$$

If $u^1(\theta_1,a)$ has ID in $(\theta_1;a_2)$, then

$$\int_{\Theta_1\times S_1} u^1_{a_2}\big(\theta_1, a_1^\star(s_1;\rho), a_2^\star(s_2;\rho)\big)d\boldsymbol{F}(\theta_1,s_1|s_2;\rho)$$

in an increasing function of s_2. From Lemma 1, responsiveness with a higher mean is equivalent to

$$\int_t^1 \Big(a_2^\star(s_2;\hat{\rho}) - a_2^\star(s_2;\rho)\Big) ds_2 \geq 0$$

for all $t \in [0,1]$. Using Lemma 6, we can then conclude that

$$\int_{S_2} \Big(a_2^\star(s_2;\hat{\rho}) - a_2^\star(s_2;\rho)\Big) \int_{\Theta_1 \times S_1} u^1_{a_2}\big(\theta_1, a_1^\star(s_1;\rho), a_2^\star(s_2;\rho)\big) d\boldsymbol{F}(\theta_1, s_1|s_2;\rho) ds_2$$

$$\geq \underbrace{\int_{S_2} \Big(a_2^\star(s_2;\hat{\rho}) - a_2^\star(s_2;\rho)\Big) ds_2}_{\geq 0 \text{ by responsiveness}} \int_{\Theta_1 \times S_1} \underbrace{u^1_{a_2}\big(\theta_1, a_1^\star(s_1;\rho), a_2^\star(0;\rho)\big)}_{\geq 0 \text{ by positive externalities}} d\boldsymbol{F}(\theta_1, s_1|0;\rho) \geq 0.$$

In either case, we get the desired result that $-VT(\rho;\hat{\rho}) \geq 0$. ■

6.4 Blackwell, Lehmann, and Monotone Information Order

It is natural to ask why the monotone information order is the relevant order to consider instead of the more familiar Blackwell informativeness (Blackwell, 1951, 1953) or the Lehmann (accuracy) order (Lehmann, 1988). The answer is two-fold: The first reason for focusing on the monotone information order has to do with the value of information in the class of decision problems we consider. Blackwell (1951, 1953) shows that all decision makers value a higher quality of information if, and only if, information quality is ranked by Blackwell informativeness. Athey and Levin (2017) show that if the class of decision problems is restricted to supermodular preferences, then a higher quality of information is valuable if, and only if, information quality is ranked by the more general monotone information order. Our results further solidify the link between the class of supermodular payoffs and the monotone information order by providing conditions on the marginal utilities of supermodular payoff functions such that, agents are more responsive when information quality increases if, and only if, information quality is ranked by the monotone information order.

Second, within the class of information structures that satisfy (A.5), the monotone information order is a more general ordering than Blackwell informativeness and the Lehmann ordering. In particular, if information structures satisfy the MLRP property (a stronger assumption than (A.5)), then Blackwell informativeness implies the Lehmann order which in turn implies the monotone information order. The converse however is not true, as shown by the example below. Figure 7 depicts the nesting of information orders and the associated class of decision problems.

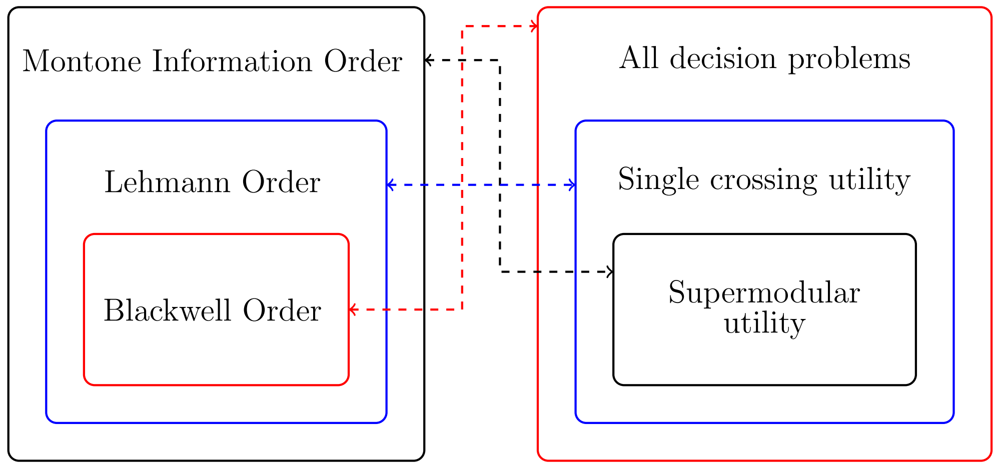


Figure 7: Information ordering and decision problems

The following is an example of information structures that can be ordered using the monotone information order but not the Lehmann order.[25] For this section only, we consider information structures $\Sigma_\rho \triangleq \big(S, \{F(\cdot|\theta;\rho)\}_{\theta\in\Theta}\big)$ such that $\{F(\cdot|\theta;\rho)\}_{\theta\in\Theta}$ satisfies the MLRP property, i.e., for any $s < s'$, the likelihood function

$$\frac{f(s'|\theta;\rho)}{f(s|\theta;\rho)}$$

is non-decreasing in θ.[26]

Lehmann (Accuracy) Order: $\Sigma_{\rho''}$ dominates $\Sigma_{\rho'}$ in the Lehmann order, denoted $\rho'' \succeq_L \rho'$, if for all $s \in S$,

$$F^{-1}\Big(F(s|\theta;\rho')\Big|\theta;\rho''\Big)$$

is non-decreasing in θ.

Example: Let $\theta \in \{\theta_1, \theta_2, \theta_3\}$ with $\theta_1 < \theta_2 < \theta_3$. Let μ_i^o be the mass at θ_i with $\mu_1^o = \mu_2^o = \frac{2}{5}$ and $\mu_3^o = \frac{1}{5}$. Consider two information structure $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$ such that the signal space S is the unit interval for both structures and $F(s|\theta_i;\rho')$ is given by

[25] See Lehmann (1988), Persico (2000), and Jewitt (2007) for a more complete analysis of the Lehmann ordering.

[26] This is a more restrictive assumption on signal structures than (A.5).

	$0 \leq s < 1/2$	$1/2 \leq s \leq 1$
θ_1	$s\frac{3}{2}$	$\frac{1+s}{2}$
θ_2	s	s
θ_3	0	$2s-1$

while $F(s|\theta_i;\rho'')$ is given by For both information structures, the marginal on the signal is

	$0 \leq s < 1/2$	$1/2 \leq s \leq 1$
θ_1	$2s$	1
θ_2	$\frac{s}{2}$	$\frac{3s-1}{2}$
θ_3	0	$2s-1$

simply the uniform distribution on $S = [0,1]$, i.e., $F_S(s;\rho') = F_s(s;\rho'') = s$ for all $s \in [0,1]$. Furthermore, both structures satisfy the MLRP property: for any $s < s' < 1/2$ or $1/2 \leq s < s'$, the likelihood functions satisfy

$$\frac{f(s'|\theta_i;\rho')}{f(s|\theta_i;\rho')} = \frac{f(s'|\theta_i;\rho'')}{f(s|\theta_i;\rho'')} = 1 \qquad \forall i = 1,2,3,$$

while for any $s < 1/2 \leq s'$, the likelihood ratios satisfy

$$\underbrace{\frac{f(s'|\theta_1;\rho')}{f(s|\theta_1;\rho')}}_{=1/3} < \underbrace{\frac{f(s'|\theta_2;\rho')}{f(s|\theta_2;\rho')}}_{=1} < \underbrace{\frac{f(s'|\theta_3;\rho')}{f(s|\theta_3;\rho')}}_{=\infty},$$

and

$$\underbrace{\frac{f(s'|\theta_1;\rho'')}{f(s|\theta_1;\rho'')}}_{=0} < \underbrace{\frac{f(s'|\theta_2;\rho'')}{f(s|\theta_2;\rho'')}}_{=3} < \underbrace{\frac{f(s'|\theta_3;\rho'')}{f(s|\theta_3;\rho'')}}_{=\infty}.$$

As a result, $s' > s$ implies $\mu(\cdot|s';\rho) \succeq_{FOSD} \mu(\cdot|s;\rho)$, for $\rho = \rho', \rho''$ (Milgrom; 1981).
We first show that $\rho' \not\succeq_L \rho''$ and $\rho'' \not\succeq_L \rho'$. If $\rho' \succeq_L \rho''$, then

$$F^{-1}\Big(F(s|\theta;\rho'')\big|\theta;\rho'\Big)$$

must be increasing in θ for every $s \in [0,1]$. However, for all $s \in [0,1]$

$$F^{-1}\Big(F(s|\theta_3;\rho'')\big|\theta_3;\rho'\Big) = s$$

whereas

$$F^{-1}\Big(F(s|\theta_1;\rho'')\big|\theta_1;\rho'\Big) \geq F^{-1}\Big(F(s|\theta_1;\rho')\big|\theta_1;\rho'\Big) = s$$

since $F(s|\theta_1;\rho'') \geq F(s|\theta_1;\rho')$. Similarly,

$$F^{-1}\Big(F(s|\theta_2;\rho'')\big|\theta_2;\rho'\Big) \leq F^{-1}\Big(F(s|\theta_2;\rho')\big|\theta_2;\rho'\Big) = s$$

because $F(s|\theta_2;\rho'') \leq F(s|\theta_2;\rho')$. Altogether, we have

$$F^{-1}\Big(F(\cdot|\theta_2;\rho'')\big|\theta_2;\rho'\Big) < F^{-1}\Big(F(\cdot|\theta_3;\rho'')\big|\theta_3;\rho'\Big) < F^{-1}\Big(F(\cdot|\theta_1;\rho'')\big|\theta_1;\rho'\Big)$$

violating the Lehmann monotonicity condition. Thus, $\rho' \not\succeq_L \rho''$.
Figure 8 depicts the conditional distributions of the signals. The solid black line is the conditional distribution of signals given θ_3 under both $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$. The solid and dashed blue lines are the conditional distribution of signals given θ_1 under $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$ respectively. Similarly, solid and dashed red lines are the conditional distribution of signals given θ_2 under $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$ respectively. Starting from $s^* \in [0,1]$, the arrows show the transformation to $\tau_i = F^{-1}\Big(F(s^*|\theta_i;\rho'')\big|\theta_i;\rho'\Big)$ where the blue, red, and black arrows correspond to θ_1, θ_2, and θ_3 respectively. Similarly, If $\rho'' \succeq_L \rho'$, then

$$F^{-1}\Big(F(s|\theta;\rho')\big|\theta;\rho''\Big)$$

must be increasing in θ for every $s \in [0,1]$. However, for all $s \in [0,1]$,

$$F^{-1}\Big(F(s|\theta_3;\rho')\big|\theta_3;\rho''\Big) = s$$

whereas

$$F^{-1}\Big(F(s|\theta_1;\rho')\big|\theta_1;\rho''\Big) \leq F^{-1}\Big(F(s|\theta_1;\rho'')\big|\theta_1;\rho''\Big) = s,$$

and

$$F^{-1}\Big(F(s|\theta_2;\rho')\big|\theta_2;\rho''\Big) \geq F^{-1}\Big(F(s|\theta_2;\rho'')\big|\theta_2;\rho''\Big) = s.$$

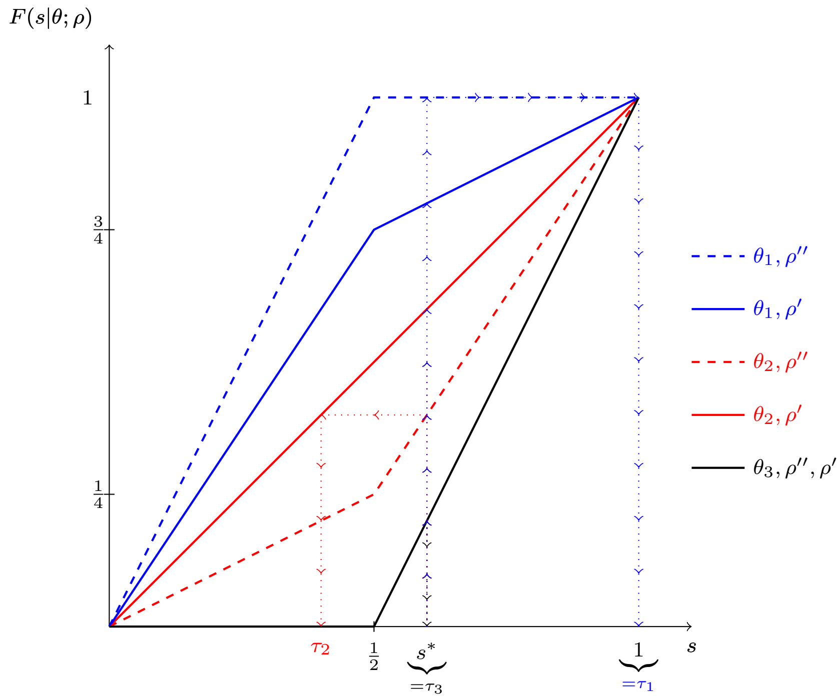


Figure 8: $\rho' \not\succeq_L \rho''$

Altogether, we have

$$F^{-1}\Big(F(\cdot|\theta_1;\rho')\big|\theta_1;\rho''\Big) < F^{-1}\Big(F(\cdot|\theta_3;\rho')\big|\theta_3;\rho''\Big) < F^{-1}\Big(F(\cdot|\theta_2;\rho')\big|\theta_2;\rho''\Big)$$

violating the Lehmann monotonicity condition. Thus, $\rho'' \not\succeq_L \rho'$. Furthermore, $\Sigma_{\rho''}$ and $\Sigma_{\rho'}$ are also not Blackwell ordered since Blackwell ordering implies Lehmann ordering (within the class of information structures with MLRP property).

Figure 9 depicts the conditional distributions of the signals. The solid black line is the conditional distribution of signals given θ_3 under both $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$. The solid and dashed blue lines are the conditional distribution of signals given θ_1 under $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$ respectively. Similarly, solid and dashed red lines are the conditional distribution of signals given θ_2 under $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$ respectively. Starting from $s^* \in [0,1]$, the arrows show the transformation to $\tilde{\tau}_i = F^{-1}\Big(F(s^*|\theta_i;\rho')\big|\theta_i;\rho''\Big)$ where the blue, red, and black arrows correspond to θ_1, θ_2, and θ_3 respectively.

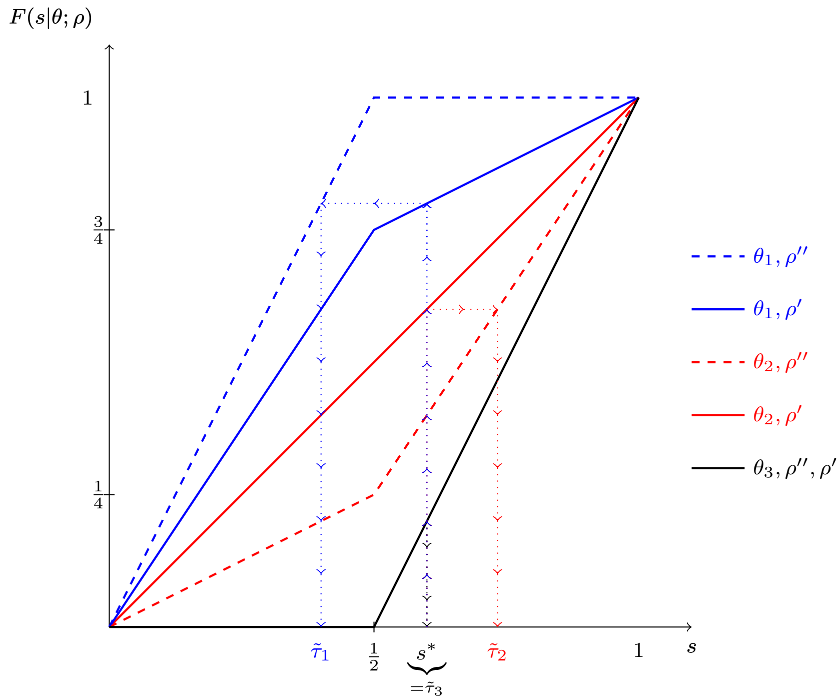


Figure 9: $\rho'' \not\succeq_L \rho'$

Next, we show that $\rho'' \succeq_{MIO} \rho'$. From Lemma 2, $\rho'' \succeq_{MIO} \rho'$ if $F(\theta_i, s; \rho') - F(\theta_i, s; \rho'') \leq 0$ for all (θ_i, s). Notice that for all $s \in [0, 1]$,

$$F(\theta_1, s; \rho') - F(\theta_1, s; \rho'') = \mu_1^o \Big(F(s|\theta_1; \rho') - F(s|\theta_1; \rho'') \Big) \leq 0.$$

Furthermore, for all $s \in [0, 1]$,

$$\begin{aligned} F(\theta_2, s; \rho') - F(\theta_2, s; \rho'') =& \mu_1^o \Big(F(s|\theta_1; \rho') - F(s|\theta_1; \rho'') \Big) + \mu_2^o \Big(F(s|\theta_2; \rho') - F(s|\theta_2; \rho'') \Big) \\ =& \frac{2}{5} \Big(F(s|\theta_1; \rho') - F(s|\theta_1; \rho'') + F(s|\theta_2; \rho') - F(s|\theta_2; \rho'') \Big) = 0. \end{aligned}$$

Finally, $F(\theta_3, s; \rho') - F(\theta_3, s; \rho'') = \sum_{i=1}^{3} \mu_i^o \Big(F(s|\theta_i; \rho') - F(s|\theta_i; \rho'') \Big) = F_S(s; \rho') - F_S(s; \rho'') = 0$. Hence, $\rho'' \succeq_{MIO} \rho'$.